AD 675992

INFORMATION SCIENCES - 1967

A Summary Report of the FY 1967 Active Contracts
and Grants of the Directorate of Information Sciences

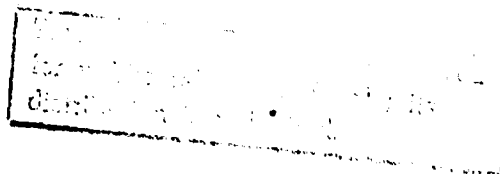**AIR FORCE OFFICE OF SCIENTIFIC RESEARCH**

Office of Aerospace Research, USAF

Arlington, Virginia 22209

D D C

OCT 16 1968

AFOSR 68-0006

PAGES_____ ARE MISSING IN ORIGINAL DOCUMENT

TABLE OF CONTENTS

PREFACE

This volume marks the tenth anniversary of the information sciences program of the Air Force Office of Scientific Research. During these ten years AFOSR has been part of three organizations; the Air Research and Development Command, the Air Force Research Division and the present Office of Aerospace Research; AFOSR has had seven Commanders or Executive Directors; the program has been funded under at least four different projects and has lived in five different boxes on the organizational chart.

This research program actually started in 1956, with the transfer of one contract with Documentation Inc from the Directorate of Advanced Studies. By 1957 the program had all of nine efforts: three for the support of publications transferred from the Mechanics Division of the Directorate of Aeronautical Sciences; two for the support of information systems transferred from the Aeromedical Division, and four of our own!

And what has happened to those original efforts?

With the help of the Office of Naval Research and a contract with the American Institute of Physics the journal, *Physics of Fluids,* is now in its 10th volume and has become completely self-sustaining.

The 16 page *International Aeronautical Abstracts,* funded through the Institute of Aeronautical Sciences, has become the giant *International Aerospace Abstracts* (occupying 9 feet of library shelves), NASA's principal method of distributing abstracts of the journal literature.

Biosciences Information Exchange, funded jointly for Defense, by AFOSR and ONR has become the Science Information Exchange, funded at the rate of $1.5 million a year by the National Science Foundation.

An effort with the Science and Technology Division of the Library of Congress to produce what was then called "Bibliography of AFOSR TN's" has produced the five volume "Air Force Scientific Research Bibliography" which gives indexed abstracts of all AFOSR sponsored publications from 1950 through 1961.

Our contract with Documentation Inc., under the leadership of the late Mortimer Taube (who took much of the zest in documentation with him when he died in September, 1965) started out as a series of theoretical papers in information storage and retrieval, scooping the gigantic International Conference on Scientific Information by at least six months and producing one book

("Information Storage and Retrieval", Columbia U Press, 1958) the title of which, at least, is still in demand today.

Then as now, the information sciences were suspect as impractical, and Taube and I were under some pressure to show that they had practical applications. We took three theoretical papers on indexing, a piece of hardware produced at no cost to the government as a byproduct of one of these papers, a stack of DD-613's and a RAMAC and built project ECHO (Experimental Contract Highlight Operation)–the first successful attempt by any Federal Agency to put its contract and grant files into computer store. And, way back in 1959, this computer store was accessible by a teletypewriter in my office!

ECHO itself smothered from a surfeit of requests for printouts–the happiest way for an information system to go–but from its ashes sprang MASIS, the management information system of the Office of Aerospace Research. And the statement has been made, by Mel Day of NASA, that ECHO served as the prototype of both the Department of Defense and the NASA mechanized contract information systems.

And, our one really long range contract, with R. Solomonoff of Zator Company to study the mechanization of inductive inference hasn't done too badly. Academedician A. N. Kolmogorov of the USSR, in his keynote paper "Foundations of Information Theory" at the International Conference on Information Theory held at San Remo, Italy this last summer cited Solomonoff's work as the foundation for his new field of algorithmic information theory which Kolomogorov claims is much broader in scope than current information theory based on probabilistic theory and entropy defined in the usual probabilistic way.

So much for 1957. What's happened since then?

Obviously the program has grown, from 9 efforts to the 81 reported in this volume. Perhaps more importantly, this growth has been qualitative as well as quantitative. Back in 1957 we were concerned with information storage and retrieval in a fairly narrow sense, and might even have been content to live under the rubric of "documentation–the theory and practice of producing storable items of information in a form of convenience to the subsequent researcher; and so organizing them as to facilitate their retrieval and the dissemination of the information they contain." (AFOSR–5266. *Glossary of Stinfo Terminology*, compiled by Dana L. Thompson. October, 1963.)

Documentation as so defined continues to be an important part of our program but the main thrust has moved towards what a committee of the Association for Computing Machinery has called

"computer science" (*Communications of the ACM 8*, 544. 1966).
They wrote:

"Computer science is concerned with *information* in much the
same sense that physics is concerned with *energy;* it is devoted
to the *representation, storage, manipulation* and *presentation*
of information in an environment permitting automatic informa-
tion systems. As physics uses energy transforming devices, com-
puter sciences uses information transforming devices. Some forms
of information have been more thoroughly studied and are better
understood than others; nevertheless, all forms of information--
numeric, alphabetic, pictorial, verbal, tactile, olfactory, results of
experimental measurement, etc.–are of interest in computer
science."

Having somewhat less of a vested interest in computers than does
the Association for Computing Machinery, and being more interested
in what computers can do than in the sullen beasts themselves,
we chose and choose to call our program *information sciences*
rather than "computer sciences". But, with the possible exception
of olfactory information, responsibility for which we are glad to
transfer elsewhere, computer sciences as so defined is now an im-
portant part of our research program.

One other change has occurred. For all practical purposes, Air
Force funding of this program has been stationary since FY 63 when,
after a rapid initial growth period, it leveled off at about the million
dollar mark. Since our last report ("Information Sciences, 1965".
AFOSR 66-0130. January, 1966), we have received massive trans-
fusions from the Advanced Research Projects Agency, through their
Directorate of Behavioral Sciences; their Directorate of Information
Processing Techniques and, in one case, through the Office of the
Director of Technical Information, Office of the Director of De-
fense Research and Engineering. The nine ARPA supported efforts
are so indicated in the text and are also listed in a separate index.

All efforts reported are those active within FY 67–that is, those
efforts in being during the period 1 July 1966 through 30 June,
1967. This differs from the three previous editions of "Information
Sciences", which reported on the calendar year rather than the
fiscal year. Because, however, of our not always successful attempts
to secure stability in support by funding efforts for more than one
year, the work reported herein includes efforts funded with FY 66,
67 and 68 funds.

The bibliographies to the separate contracts and grants list 13 books,
96 book chapters (the majority of these being papers included in
symposium proceedings), 134 journal articles and 171 reports.
(This tabulation does not include the publications from the ARPA
sponsored effort at Carnegie Mellon University, since this single
effort is as large as the rest of the program.) Every effort has been

made to obtain either AFOSR or AD numbers for these publications. Qualified requestors may obtain copies of these, numbered, publications from the Defense Documentation Center, Alexandria, Virginia 22314. Those not eligible for DDC services may buy the publications, for $3.00 each, from the Clearinghouse, US Department of Commerce, Springfield, Virginia 22151.

The project descriptions were written by the Information Science project monitors, Rowena W. Swanson and Captain Eliot Sohmer, and reviewed by the principal investigators with a necessary minimum of editorial revision. The manuscript was edited by Captain Sohmer and typed by Mrs. Loretta Burke.

It is rather startling to me, looking backwards over the 10 years that this research program has been in existence, to realize how few people nave been involved with it. First, and of course foremost, my secretaries: Mrs. Elsie Reed (then Mrs. Glover); Mrs Melver Hodgson and my current nonpareil, Mrs. Texie Lee Gleeson. Major (then Captain) Joseph Martino worked half time on the program during 1960 and 1961.

Mrs. Rowena W. Swanson came to AFOSR in 1961, from the Office of Research and Development, U.S. Patent Office and has assumed greater and greater responsibility, including serving as Acting Director during the year of 1965-1966 while I was seconded to the Office of the Director of Technical Information, DDR&E. Captain Thomas K. Burgess came aboard in April of 1964 and left in December, 1965 to begin a civilian career by mechanizing the library at Washington State University, Pullman, Washington. His replacement, Captain Eliot Sohmer, joined us in October of 1965. And, somehow, almost 12 years have gone by since I first came to AFOSR.

HAROLD WOOSTER   Arlington, Virginia
Director of Information Sciences   January, 1968

viii

PUBLICATIONS OF THE DIRECTORATE OF
INFORMATION SCIENCES · 1965-1967

Eliot Sohmer, "The Computer and Its Languages", in *AFOSR Research* (Government Printing Office, 1967; AFOSR 67-0300, AD 659366) pp. 219-222.

Rowena W. Swanson, "Move the Information...A Kind of Missionary Spirit" (Government Printing Office, June 1967; AFOSR 67-1247, AD 657794).

Rowena W. Swanson, "Influences from Cybernetics on Information Sciences", in AFOSR Research (Government Printing Office, July 1967; AFOSR 67-0300, AD 659366) pp. 217-219.

Harold Wooster, *Communications of the ACM*, "Centralized Information Services, A Panel Discussion", pp. 708-709 (November, 1965).

Harold Wooster, *Journal of the Association of Official Analytical Chemists*, "Whither Information", 49, pp 93-99 (1966).

Harold Wooster, *IEEE Spectrum*, "Communication with Extra-Terrestrial Intelligence", pp. 153-163 (March 1966).

Harold Wooster, *Computers and Automation*, "The Quest and the Covenant--How to Translate Ideas Into Government Research Contracts", pp. 2-8 (May 1967).

Harold Wooster, *Library Journal*, "Books and Libraries in the Scientific Age", pp. 2511-2515 (July 1967).

Harold Wooster, "The Zoo and the Jungle--A Comparison of the Information Practices of Intelligence Analysts and of Scientists", in *Information System Science and Technology*, Donald Walker, Ed. (Thompson Book Company, Washington, D.C., 1967) pp. 307-315.

Harold Wooster, "Policy Planning for Technical Information in Industry", to appear in symposium proceedings, Bad Godesberg, Germany, November 1967; AFOSR 67-2588.

Harold Wooster, *OAR Research Review*, "Basic Research and The Department of Defense", pp. 1-9 (October 1967).

Section 1

INFORMATION SYSTEMS RESEARCH

Information systems research looks at the structure and operation of entire systems or of units within a system viewed in the framework of the entire system. Efficiency and effectiveness of methods and tools for the input, throughput, dissemination, and use of information are principal topics for investigation.

Information systems abound. They are used for document control, data control, logistics, management and command, and intelligence within military and civilian contexts. Many existing systems are based on *ad hoc* planning and have suffered explosive, unanticipated growth. Frequently mechanization or other changes have been introduced to alleviate some information processing difficulties, but the effects are usually far from optimal.

Questions must be asked about systems in the abstract. What are possible information flow patterns? With what configurations of men and machines can each be achieved? What is the match between pattern and purpose of the system? What effect does the environment have on the system? Who feeds the system? Who uses it? Who could feed it or use it? What are the tradeoffs for a given system for a given purpose?

Questions must be asked about existing systems. Existing structures and operations must be quantitatively and realistically analyzed. The effects of perturbations must be examined in the abstract and tested in simulations and in ongoing operations. Live data under perturbations are invaluable but difficult to obtain.

1

1.1 DEVELOPMENT OF A STANDARDIZED LANGUAGE FOR DESCRIBING DOCUMENTATION SYSTEMS

Lea M. Bohnert

CEIR Inc,, Arlington, Va.

Project Description

A profusion of documentation systems exists today in military and other government establishments for recording and processing data and information. Most systems were designed and are being operated by nondocumentation specialists, since documentation as a discipline is relatively new. This effort is for the analysis of the fundamental operations common to all documentation systems for the purpose of standardizing the description of these operations (e. g., routing, summarizing, storing, and searching) as an aid to the intelligent modification of existing systems and the design of new systems. A checklist will be developed to be used in the evaluation of existing documentation systems, including comparisons of structure, operations, and costs. The applicability of computer-based display techniques to documentation operations will be examined toward the development of procedures for integrating machines into documentation systems for various user needs. The study should provide a systematic approach to the analysis and evaluation of documentation systems which can be applied in cost and quality control, and indications of man-machine combinations which take advantage of the volume and repetitive processing capabilities of machines.

Summary of Recent Research Results

The aim of this research effort is to distinguish and describe retrieval activities that are being performed within any organization with a library type mission. In order to develop a standardized method of describing and analyzing retrieval activities, it is necessary to uniquely define what one means by "retrieval work". A model was developed based on the following assumptions:

1. Retrieval is defined as a documentation, or message delivery, service that provides documents only upon requests by readers.

2. Only two functions are performed by a retrieval service. Upon request, the service must *identify* and *provide* appropriate documents, if any are available.

2

3. Robert A. Fairthorne's "marking and parking" terminology is used to analyze and describe the visible work performed by a retrieval service in an attempt to distinguish that which may be delegatable to clerks or machines.

Three main results have been obtained:

1. The physical objects of main interest in retrieval have been classified into four records:

   a. The document;

   b. document substitutes (any record a reader will accept in lieu of the original);

   c. retrieval aids (records designed to be manipulated instead of the original document e.g. charge-out slips, catalog cards, etc...);

   d. data (documents composed of items of equal interest to the reader).

2. The methods of manipulating records have been classified into five operations:

   a. preparation;

   b. matching;

   c. transport;

   d. storage; and

   e. reproduction.

Retrieval processes such as acquisition, indexing, current dissemination, and searching, have been shown to be combinations of these five operations applied to the four types of retrieval records.

3. Document requests have been classified into two parts:

   a. A description of the requested document; and

   b. A specification of the types of physical access to the document.

It is anticipated that this study will lead to the publishing of a textbook for library school graduate students, interested laymen, and as a basis for further research.

Reports and Publications

Lea M. Bohnert, "Retrieval of Technical Documents"
   (Final Technical Report, November 1967).

3

1.2 DIFFUSION OF ABSTRACTING SERVICE MEDIA FROM GOVERNMENT—SPONSORED RESEARCH

Irving M. Klempner

Columbia University, New York, New York

Project Description

Reports of federally sponsored activities in science and technology are documented by various information services, and information is disseminated by such abstract bulletins as *Technical Abstracts Bulletin* (DOD), *Scientific and Technical Aerospace Reports* (NASA), and *Nuclear Science Abstracts* (AEC). The aim of this effort is to determine whether these media are reaching the populations for whom they are intended. Patterns of dissemination and use will be examined both in defense-oriented industries and in industries and institutions having few or no government-sponsored activities. Data will be collected by questionnaire and interview methods. Analyses will be made according to the nature of the user and by geographic distribution. Results should aid in assessing the potential of abstracting service media for the diffusion of research results and in formulating policies and improving procedures for the dissemination of information.

Summary of Recent Research Results

This effort examined the use and nonuse patterns for 4 federal index-and-abstract services: AEC's *Nuclear Science Abstracts (NSA)*; NASA's *Scientific and Technical Aerospace Reports (STAR)*; DDC's *TAB*, and the Dept. of Commerce's *U.S. Government Research and Development Reports (USGRDR)*. A culling of mailing lists of subscribers to these services disclosed 7949 non-federal U.S. recipients of one or more copies of the services and a list of combinations of services received at a total of 4179 sites (3004 discrete organizations). Almost 50% of the recipients are located in 5 states: California, New York, Massachusetts, Pennsylvania, and New Jersey. Of 1139 questionnaires (12 pages, 50 questions) mailed to a stratified random sample, 823 were returned (776 usable). Analysis of the data disclosed:

- 82% of the host organizations engaged in federal R&D work, about 50% employing 300 or more scientists and engineers; 96% maintained a library or information center, but about 50% employed one or less professional librarian

4

- among the respondents, 47% were acting as librarians, 22% were engaged in R&D, and 20% were managers or administrators
- almost all respondents held a bachelor's degree; 25% held Ph. D.'s; about 40% are estimated in the 30-40 age group
- 42% had task assignments at least once annually outside their field of specialization
- 91% used an index-and-abstract service over a specified 6-month period; regular use was indicated by 60% to acquire specific information directly related to an on-going project (57.1%), for current awareness in a primary field of interest (42.2%), and for current awareness in a secondary field of interest (45.3%)
- 63.9% preferred present coverage to a more specialized subject service, 47.1% being satisfied with the present format
- about 30% of the libraries do not permit circulation of the services; the collections of 40% exceed 15,000 report titles (15% exceed 100,000 titles)
- of recipient libraries, 80.2% of the industrials and 19.8% of the academics catalog technical reports (40% assigning 5 or more descriptors); 83.5% and 16.5% prepare abstracts, respectively; 83.5% and 16.5% issue acquisition bulletins, respectively; and 85.2% and 14.8% maintain an SDI service, respectively
- more than 50% of the libraries retain copies of the services for 3 or more years
- 82.3% expressed no difficulty in acquiring announced publications

The following data was used to isolate research-oriented non-recipients of the services: 3260 industrial firms listed by *Industrial Research Laboratories of the United States* (1965) (believed to represent more than 95% of the dollar volume of the U.S. industrial R&D effort); 1096 institutions listed in the *Roster of U.S. Government Research and Development Contracts in Aerospace and Defense* (awarded 7500 contracts in FY 64); 2775 firms on the *List of Small Business Concerns Interested in Performing Research and Development* (1963); and 3188 organizations in the *Research Centers Directory* (1965) (3014 located in the U.S.). Checks against his recipient file disclosed that 68.6%, 34.1%, 86.4%, and 92% (of 3014), respectively, were non-recipients of any of the 4 services. He mailed non-recipient questionnaires to a random sample of 480 and received 235 replies. His analysis shows:
- respondents primarily in electronics and engineering (24.4%), chemistry and chemical engineering (16.2%), and other engineering (15.8%)
- 81.3% of the host organizations employ fewer than 20 scientists and engineers; 87% conduct in-house research (about 50% of this for private organizations and 18% for the federal government)
- 75% maintain libraries, 12% of these being staffed by professional librarians
- of the individuals who replied, 63% were in management, 27% were engaged in R&D, and 1.7% were librarians or information services personnel

5

• 48% of the respondents believed their information needs were being met "fairly well" and 13% "inadequately;" 63% needed information outside their field of specialization during a specified 12-month period

• 23% were *Chemical Abstracts* subscribers; less than 10% received any other well-recognized service; 86% had no knowledge of *NSA* or *TAB*; 77% and 71% did not know about *STAR* and *USGRDR*, respectively; and 60% wanted information about one or more of these services

• most respondents favored an index-and-abstract service, 26% within their specialty and 54% in a broader primary field; 53% most often needed specific data for assigned tasks, 30% also needing information about laboratory techniques and apparatus.

Reports and Publications

Irving M. Klempner, "Diffusion of Abstracting and Indexing Services for Government-Sponsored Research". Scarecrow Press, Metuchen, New Jersey, 1968

6

1.3 ANALYSIS OF INFORMATION SCIENCE TERMINOLOGY

Charles W. Shilling and Patricia Fuellhart

George Washington University, Washington, D.C.

Project Description

The field of information sciences is relatively new. As is characteristic of interdisciplinary fields there is little agreement as to the boundaries of the field, less agreement as to what documents are relevant, and perhaps even less agreement as to the terms used to designate these documents. There is inevitable duplication of effort in the development of information science vocabularies since individual researchers are often unaware of similar activities. The first step of this research will be to search out, on an international basis, individuals and organizations developing word lists, glossaries, dictionaries, subject category headings, thesauri and other authority files in the field of information science and documentation and publish a report of individuals and organizations so engaged. The vocabularies being compiled by these groups will be compared and analyzed to determine terms common to all lists, terms with a lesser degree of overlap, and terms that are discrete and apparently essential to the specific lists but with no significant overlap between lists. A consolidation of these lists may be attempted.

Summary of Recent Research Results

This AFOSR research effort received its initial sponsorship less than one year ago, and will be reported in the next issue of this publication.

1.4   OPTIMIZATION AND EVALUATION OF MEMORY ORGANIZATIONS FOR LARGE SCALE FILING SYSTEMS

Kenneth B. Krohn

Krohn Rhodes Research Institute, Inc., Washington, D.C.

Project Description

The problem of file handling is often viewed as the last great software problem to make large-scale time-sharing systems operationally effective. If large files cannot be effectively managed, i.e., data stored and retrieved quickly, the system is worthless. Several solutions exist, but none is entirely satisfactory with respect to speed, cost, and organizational structure. This effort will investigate methods of analysis and simulation of projected memory-filing systems to evaluate their effectiveness in such a way as to indicate their potential for data retrieval and other applications, both by developing theoretical means for determining projected efficiencies and by enabling practical testing and comparison of existing systems. Basic theorems in the theory of the class of finite semigroups that characterize combinatorial machines will be applied in conjunction with simulation techniques to develop a method of comparing system efficiency with its theoretical maximum. A preliminary determination of a repertory of schematic data bases will be made that will enable practical configurations to be tested, evaluated, and modified to maximize efficiencies in given contexts.

Summary of Recent Research Results

This AFOSR research effort received its initial sponsorship less than one year ago, and will be reported in the next issue of this publication.

8

1.5   # STUDY OF THE MAN-SYSTEM
INTERFACE IN LIBRARIES
RETRIEVAL STUDIES

Robert S. Taylor

Lehigh University Bethlehem, Pennsylvania

Project Description

Concern with user-oriented services of a library and systematization
and improvement of the user interface with a store of records to opti-
mize his retrieval and use of the information therein has come under
scientific investigation only recently. Studies of applications of ma-
chine technology to this problem area are usually not from the view-
point of the real user and potentialities of improvement in on-going
library systems. This effort postulates the library as an evolutionary
information system and proposes the analysis of the various user-
library interface points toward their explication with a view toward
merging productive aspects of computer and systems sciences knowl-
edge with results of library system studies. Parameters of the man-
library system interface processes will be identified qualitatively
and quantitatively where possible. Restructuring these interfaces
through the introduction of methods and devices will be examined
and evaluated. Among the goals is a better understanding of the
processes involved in "negotiating the question" and the establish-
ment of implementable procedures for effective interfacing in on-
going systems.

Summary of Recent Research Results

A study of the information-seeking behavior of professional
personnel revealed that ease of use; i. e., access, is more important
than amount or quality of information. A study of question
negotiation isolated five filters consciously sought by librarians
when they negotiate an inquiry. These are (1) subject definition;
(2) objective and motivation; (3) personal characteristics of the in-
quirer: (4) relationship of inquiry to file organization; and (5) an-
ticipated or acceptable answers.

9

An on-line computer program for negotiating inquiries (GRINS) was developed, in which a prediction algorithm is provided to determine the rate at which the definition of the inquirer's need develops. The program interrogates the user in a natural language conversational mode. When the negotiation is judged by GRINS to be as well developed as it is likely to get, a search is made of the available documents. An ordered list is produced which comes closest to the user's expressed need.

Reports and Publications

C.E. Hieber, "An Analysis of Questions and Answers in Libraries" (Studies in the Man-System Interface in Libraries (SMSIL) Report Number 1, June 1966; AD 635020).

D.J. Hillman, "Computational, Phonological and Morphological Linguistics and Retrieval Studies" (Grammars and Text Analysis Report Number 1, 1965, AFOSR 65-1892, AD621128).

J.S. Green, "GRINS, An On-Line Structure for the Negotiation of Inquiries,"(SMSIL Report Number 4, September 1967).

V. Rosenberg, "The Application of Psychometric Techniques to Determine the Attitudes of Individuals Toward Information-Seeking" (SMSIL Report Number 2, July 1966; AD 637713).

R.S. Taylor, "Question Negotiation and Information-Seeking in Libraries" (SMSIL Report Number 3, July 1967).

R.S. Taylor, R.R. Anderson, A.F. Amico and J.S. Green, "Systems Manual for the Experimental Literature Collection and Reference Retrieval System of the Center for the Information Sciences" (ERSS Report Number 2, 1967; AFOSR 67-1676, AD 653279).

PRINCIPLES AND IMPLEMENTATION TECHNIQUES FOR INFORMATION SYSTEM DESIGN

Morris Rubinoff

University of Pennsylvania Philadelphia, Pennsylvania

Project Description

The need for information in specialized subject fields has often resulted in the ad hoc establishment of storage and retrieval systems which are not user oriented and, hence, seldom used. This effort supports the development of technical principles and implementation criteria for mechanized information storage and retrieval systems that will hopefully result in better systems than those based on the ad hoc approach. The study will concentrate on four major areas at the system level-

1. Specifications for "Real English" as an input query language for the system user. This is an advanced version of "Easy English", the presently used language, with stronger linguistic foundations possessing firmer roots in formal English grammar - 2. derivation of objective measures for terms that convey subjective meaning to information retrieval experts, including such expressions as "index term", "synonym", "scope", and "common term", as well as the derivation of advanced tools for making objective measurements, based on word frequencies and associations in selected texts and word inferences - 3. exploration of potential uses of monographs and state-of-the-art reviews as tools to augment or replace index terms in file search - and 4. attempt to find means for simplifying the maintenance of a mechanized information file and its associated operating system. Research to date has resulted in a well-documented and efficient retrieval system that provides much man-machine interaction. Continued study will be made of remote retrieval using on-line search with much man-machine interaction, as well as the use of graphic equipment to facilitate the search strategies of a user.

Summary of Recent Research Results

Progress has been made in two problem areas associated with modes of natural language console inquiry: 1) the semantic problems of command language for accessing the machine, and 2) the semantic problems of data/descriptor interpretation by both man and machine.

The relatively poor results obtained using Symbolic Command Language as a search language at the typewriter console led to the implementation of *Easy English*. The latter is a natural command language made up of readily recognized sentences of the English language, sentences that any layman might be expected to use in everyday requests for services or articles from a familiar source. The effectiveness of Easy English has led to its extension to *Real English*, a subset of natural English capable of growing indefinitely to serve the user of a typewriter console in every phase of man/machine communication. It is based upon a syntax analysis grammar that has been developed by Professor Naomi Sager (originally of the University of Pennsylvania, now of New York University) over a period of ten years. Real English will gradually incorporate more and more of Professor Sager's syntax analysis grammar to implement search, loading, thesaurus development, quality assurance, and all aspects of information system activities.

The data base has been augmented through inclusion of semantic tools and data/descriptor interpretation. Objective measures have been derived for various classes of terms. For example, "common terms" have been defined to be words that appear with highest frequency in the general literature. A sample of approximately 4,000,000 words was donated to the Moore School Information System Laboratory by *Time* magazine. This sample contained every word published in *Time* over a period of about one year. The most frequently appearing words (the top 1/3) were excellent representatives of common terms. These words have been adopted by the Information System Laboratory as its current list of common terms.

A more specialized text consisting of 217 reviews of articles in the computer programming field were drawn from the ACM Computing Reviews for 1962, 1964, and 1966. After common terms were deleted, the words of most frequent occurrence were obtained by computer processing. These words were separated by statistical measures into two categories: core terms and particular terms. The former were expected to be good representatives of the scope or central subject matter, the latter were expected to be excellent substitutes for index terms. Both expectations appear to have been realized, as indicated by our research on "adaptive interface",

Adaptive interface refers to a process whereby the information system gradually infers the interest profile of a searcher. The adaptive behavior is exhibited in two forms. The first form is the ability of the system to respond to an index set proposal in a manner that depends upon the previous history of the man/machine dialogue. The system does this by preparing a profile of interest during the course of the interaction and using this interest profile to frame its replies. The profile consists of numbers or scores that are related to the level of interest that the user has developed in each of a set of microthesauri. The second form of adaptive

12

behavior is the ability of the system to render suggestions to the user and to direct the interaction in accordance with its estimate of user interest profile.

The system can determine at least three types of directions in any interaction:

(1) A tendency for the profile of interest to centralize on a small group of microthesauri–in this case the user is probably refining his search in one area and should be apprised that he is reaching a terminal state in his interaction.

(2) No apparent trend in the pattern of the search–perhaps the user is browsing or he does not understand the function of the system.

(3) A tendency toward an "oscillation" in the interest profile–perhaps the user has been diverted in his search or he has changed the emphasis intentionally.

The action taken by the system ranges from advisory guidance to pointed suggestions, and ultimately to the suggestion that human intervention is required.

This system has been implemented on the 7040-PDP8 computer complex with remote teletypewriter input. The 7040 portion of the system is programmed in L^6 linked list programming language. All functions needed to implement the interaction has been included. The adaptive algorithm has been partially implemented and full implementation requires only the gradual accumulation of relation tables for a complete data base.

The techniques for objectively deriving common terms, core terms, particular terms, and two dimensional relation tables played a significant part in the successful demonstration of adaptive interface. The same techniques are now being extended to derive synonym sets and other binary word relationships by computer processing.

Reports and Publications

H. Cautin and F. Rapp, "Description of Easy English" (Moore School Report, February 1967).

H. Cautin, T.C. Lowe, F. Rapp and M. Rubinoff, "An Experimental On-Line Information Retrieval System" (Moore School Report, May 1967; also submitted to *Datamation* for publication).

John S. Edwards, "Adaptive Man-Machine Interaction in Information Retrieval (Ph. D. Dissertation, to be published).

Stephen B. Fischer, "An Executive Control System for Information Retrieval via a Remote Console" (Masters Thesis, 1966).

M. Fogel, "On-Line Typewriter Access to Classification Tables on Drum Storage" (Moore School Report, April 1967).

Thomas C. Lowe, "Design Principles for an On-Line Information Retrieval System" (Moore School Report, December 1966; AFOSR 67-0423, AD647196).

Thomas C. Lowe, "Design Principles for an On-Line Information Retrieval System (Ph. D. Dissertation, 1966).

Thomas C. Lowe, *Software Age*, "Direct-Access Memory Retrieval Using Truncated Record Names", **1**, Number 1 (September 1967).

Elayne R. Rubinoff, "Factors Involved in Testing Information Systems" (Moore School Report, August 1966).

Morris Rubinoff, *IEEE Trans. Eng. Writing and Speech*, "A Rapid Procedure for Launching a Microthesaurus" (August 1964).

Morris Rubinoff and J.F. White Jr., "Description of Cataloging and Indexing System for the ACM Repository (Moore School Report, January 1965; AFOSR 65-0424, AD612277).

Morris Rubinoff and J.F. White Jr., *Communications of the ACM*, "Establishment of the ACM Repository and Principles of the IR System Applied to its Operation", **8**, number 10, pp 595-601 (October 1965; AFOSR 66-0011, AD632185).

Morris Rubinoff, et al, "Summary Description of Easy English" (Moore School Report, February 1967).

Morris Rubinoff, S. Bergman, H. Cautin and F. Rapp, "Easy English, A Language for Information Retrieval Through a Remote Type-writer Console" (Moore School Report, April 1967; also submitted for publication to *Communications of the ACM*).

Morris Rubinoff, et al, "The Moore School Information Systems Laboratory (Moore School Report, May 1967; AFOSR 67-1952, AD657809).

Morris Rubinoff and D.C. Stone, "Semantic Tools in Information Retrieval" (Moore School Report, May 1967).

Morris Rubinoff, S. Bergman, W. Franks and E.R. Rubinoff, "Experimental Evaluation of Information Retrieval Through a Tele-typew ter" (Moore School Report, June 1967; also submitted for publication to *Communications of the ACM*).

Morris Rubinoff, W. Franks, and D. C. Stone, "Description of an Experiment Investigating Term Relationships as Interpreted by Humans (Moore School Report, June 1967).

Morris Rubinoff, *ACM Sigir Forum*, "Why Sigir", **4**, Number 2 (1967).

Morris Rubinoff and D.C. Stone, "Semantic Tools in Information Retrieval", to appear in *Proceedings of the National Convention of ADI*, 1967.

J.M. Smith, "An Oral Experiment on Retrieval Dialogue" (Moore School Report, June 1967).

Don C.Stone, "Word Association Experiments...Basic Considerations (Moore School Report, August 1966).

Don C. Stone, "Word Statistics as an Aid in Establishing Information Systems" (Masters Thesis, to be published).

14

Invited Talks

Morris Rubinoff, "Mechanized Information Retrieval", presented
    at the Institute for Applied Mathematics, The Institute Linguistics,
    The Workshop on Automata, Prague, Czechoslovakia, and the
    Institute for Computing Machinery, London England (Four papers).
Morris Rubinoff, "Computer Manipulation of Text and Pictorial
    Data", presented to the Eastman Kodak Company, September
    1967.
Morris Rubinoff, "Data Processing and System Engineering", to be
    presented at the First World Congress of Engineers and Architects
    in Israel, December 1967).

SCIENTIFIC BASES FOR
INFORMATION PROCESSING

Ralph R. Shaw

Rutgers University New Brunswick, New Jersey

Project Description

There is a need for both comparisons and evaluations of particular information systems and indexing techniques, and also for studies of basic principles pertaining to storage, retrieval and indexing operations. This effort will explore several problem areas in depth to provide tools useful for both present systems and the design of future systems. Previous studies have evaluated such areas as automatic book indexing by computer, the scope and operating efficiencies of information centers, book vs. card catalog cost, and a comparative study of three systems of information retrieval. Research is in progress on the technical qualifications required of information center personnel and on the usefulness of citation indexing. Other studies are planned on automatic indexing systems, information handling systems, quality and consistency of subject indexing and man-machine relationships in documentation processes.

Summary of Research Results

A comparative study of technical information centers in the United States and India is now in progress with a published report expected; work on the requirements and qualifications of special library and technical information center personnel is in final phase.

Mr. Theodore Shih-Shu Huang completed a study of the effectiveness of citation indexing. Three copies of current research interest were submitted by three geneticists, one each. A backward search by checking documents was done for one topic. A forward search in a citation index and a search in subject indexes were also done for all three topics. For the pertinence judgment of references retrieved by searching the citation index and the subject indexes, the retrieved documents were read in entirety by the scientist who submitted the search topic. Pertinence was defined as "the presence in the document of information that is within the scope of the topic and is also useful to research on the topic." The scope of each topic was defined by the scientist prior to the searches.

16

Two hypotheses were formulated before the data were collected. The backward search by checking documents affirms the first hypothesis that a backward search, without the elimination of noise at each step, will retrieve an overwhelming number of references, most of them being noise, and with the noise increasing at each step. The results of the subject index searches for all three topics have a pertinence ratio of 68.00%, whereas the citation index search results have a pertinence ratio of 37.50%. A statistical test affirms the second hypothesis that the difference in the proportions of pertinent references among the unique references retrieved by the two types of index searches is significant.

It was found on the basis of the investigation that the subject index search has greater efficacy than the citation index search, since the former scored better with regard to all the measures except redundancy ratio and unit search time.

Thus, the citation index search can be a supplementary aid to the subject index search when additional pertinent references are needed.

The claim that the citation index gives one the "ease with which one can obtain a negative result with a high degree of certainty" has not been borne out by the data in this research. Nor has the other claim been borne out that a citation index search retrieves pertinent references unexpectedly from fields unfamiliar to the scientist who initiates the literature search.

A backward search by checking documents leads to an immense number of references, whereas a forward search leads to a relatively small number of references.

The pertinence ratio decreases while the noise ratio increases with each step of a forward or a backward search. Cycling, therefore, does not appear to be a very promising reference retrieval aid. The ever-increasing noise indicates that it would be useless to do machine searches, either forward or backward, beyond the first step without the elimination of noise at each step. The elimination of noise at each step requires human intervention.

Reports and Publications

S. Artandi and T. C. Hines, *Amer. Doc.*, "Roles and Links - Or Forward to Cutter", **14**, pp. 74-77 (January, 1963, AFOSRJ 769).
S. Artandi, *Special Libraries*, "Special Library Services - Current Thinking and Future Trends", **54**, No. 2, pp. 103-106 (February 1963, AFOSRJ 914).
S. Artandi, "Thesaurus Controls Automatic Book Indexing By Computer", in *Proceedings 26th Annual Meeting Am. Doc. Inst.*, Chicago, Ill., October 6-11, 1963, pp. 1-2 (AFOSR 64-0235).

S. Artandi, *Library Resources and Technical Services,* "Measure of Indexing", **8**, No. 3, pp. 229-236 (Summer, 1964, AFOSR 64-2260).

S. Artandi, *Am. Doc.,* "Automatic Book Indexing by Computer", **15**, No. 4, pp. 250-257 (October, 1964, AFOSR 65-0290. AD 611-464).

R.F. Clarke, *Library Journal,* "The Impact of Photocopying on Scholarly Publishing", **88**, No. 13, pp. 2625-2629 (July, 1963, AFOSRJ 1148).

F. Heinritz, *Library Resources and Technical Services,* "Book Versus Card Catalog Costs", **7**, pp. 229-236 (1963, AFOSRJ 827).

R. R. Shaw, *The New York Times Literary Supplement,* "Electronic Storage and Searching - Freeing the Mind", pp. 19-23 (April 6, 1962, AFOSRJ 547).

R R. Shaw, *Research Management,* "The Function of a Modern Special Library", **5**, pp. 485-492 (1962, AFOSRJ 828).

R.R. Shaw, *Science,* "Information Retrieval", **140**, pp. 606-609 (May 10, 1963, AFOSRJ 493).

1.8   **DATA COLLECTION REDUCTION
ANALYSIS AND DISSEMINATION**

DeWitt O. Myatt and B.K. Farris

Science Communication Inc., Washington, D.C.

Project Description

This effort is being procured and monitored for the Task Group on
National Systems, Committee on Scientific and Technical Information
(COSATI), Federal Council on Science and Technology. This is a
systems study of national patterns of technical data handling,
with emphasis on 3 areas: data acquired in the course of experiments
and observation of natural phenomena; data such as manufacturers
catalogs, mil-specs and the like (probably the most impor..nt class
used by technical personnel); and data used in training of skilled
and semi-skilled personnel (DOD spends 3 billion a year on such
training). The study will establish how the various types and forms
of data are acquired, stored, retrieved, packaged and disseminated
for various specific types of users, why these packaging methods
have been adopted, and what changes in storage, retrieval, packaging
or dissemination are foreseen in the near future. A preliminary census
of data efforts in industry, the professions and government will be
developed to guide the formation of national policy with respect to
data collection, reduction, storage, retrieval, analysis and dissemination.

Summary of Recent Research Results

Project efforts to date reveal that the concept of national data
systems has not been previously considered by the majority of in-
dividuals and organizations involved in scientific and technical data
activities. Consequently, the study has, by necessity, included effort
toward structuring concepts of scientific and technical data, data
management, and data systems. Field survey activity has tested the
validity of these structuring concepts and has defined the types and
levels of discourse and analysis currently feasible concerning na-
tional scientific and technical data systems.

An integrated array of survey methods, including literature screening, personal interviews, workshops, and mail questionnaires, has been employed to assemble information for a census of data activities of national importance and to document current issues or problems in data activities. More than 200 formal data efforts (data collection networks, data centers, data document depositories, etc.) have been identified and statistics collected on their organization and operations. More than 200 major issues or problems, which will be considered in the development of a time-phased plan for study of national data systems concepts, have been documented. Research to collect information for current status reports describing the data activities in ten selected fields of science and technology has been conducted, and these reports are being drafted. A total of 150 personal interviews and ten workshops have been conducted. In addition, 800 mail questionnaires have been distributed to obtain information concerning data efforts and issues.

Reports and Publications

DeWitt O. Myatt and B.K. Farris, "Study of Scientific and Technical Data Activities of National Importance, Semi-annual Progress Report" (Science Communication, Inc., 28 Feb. 1967).
DeWitt O. Myatt and B. K. Farris, "Study of Scientific and Technical Data Activities of National Importance, Semi-annual Progress Report" (Science Communication, Inc., 31 Aug. 1967).

APPLICATIONS OF A MODEL FOR COMMUNICATION INFORMATION TRANSMISSION PROCESSES

William Goffman

Western Reserve University, Cleveland, Ohio

Project Description

It has been hypothesized that ability to model the communication or information transmission process would enable analysis of the dissemination of information and thereby aid in future control of such processes, both to utilize existing information and to identify areas in which additional information or refinement of existing information is needed. In this effort, a model, based on concepts of epidemic processes, has been developed. This research seeks to apply the model to the evolution of concepts in the field of symbolic logic as these have become manifest through the literature since Boole's 1847 paper. The effort will exploit Church's bibliography and several up-datings in the literature. Applications of the model will trace both total populations and groups within the total population (algebraic logic, logical foundations of mathematics, intuitive metamathematics, formalized axiomatics). Patterns in the data will be examined for their information about the transmission of ideas and their import with respect to the efficient structuring of information storage and retrieval systems.

Summary of Recent Research Results

This effort attempts to locate the fundamental properties that all communication processes have in common. Study has led to the formulation of a general mathematical theory of epidemics in which all communication processes are considered as epidemic processes. During the past year, emphasis has been placed on solving three special problems relating to epidemic theory. These are:

(1) The problem of stability of epidemic processes

(2) The problem of control of such processes

(3) The relation between co-occurring epidemics.

21

Research has established that the stability of an epidemic process is equivalent to stability in the sense of Lyapunov. Hence, it is possible to determine stable or unstable states of a communication process by applying the direct method of Lyapunov. Thus, to study a complex communication process one need not obtain the solution to the system of non-linear differential equations characterizing the process (the solution being in general not possible).

The mathematical properties required for the co-occurrence of epidemics, and the problem of evaluating the performance of an information retrieval system, have also been investigated. Since the performance of information retrieval systems can provide a control over the outcome of a communication process, it follows that the answer to the optimal control of such processes may be found by a proper evaluation of the performances of information retrieval systems. The entire theory has been expounded in (13). An experimental demonstration of the application of this theory to the transmission of ideas has been successfully carried out. The results of this experiment appear in (12).

Reports and Publications

William Goffman, *Information Storage and Retrieval*, "A Searching Procedure for Information Retrieval", 2, Number 2, pp. 73-78 (1964) AFOSR 64-2262, AD452528).

William Goffman and Vaun A. Newill, *Nature*, "Generalization of Epidemic Theory, An Application to the Transmission of Ideas", 204, Number 4955, pp. 225-228 (October, 1964; AFOSR 65-0588, AD614744).

William Goffman, *Information Storage and Retrieval*, "On Relevance as a Measure", 2, pp. 201-203 (1964; AFOSR 65-0589, AD615046).

William Goffman, J. Verhoeff and J. Belzer, *Am. Doc.*, "Use of Meta-Language in Information Retrieval Systems", 15, Number 1, pp. 14-22 (January 1964; AFOSR 65-1520, AD625577).

William Goffman and Vaun A. Newill, *Parameters of Information Science*, "Searching Titles by Man, Machine and Chance, 1, from the *Proceedings* of the Annual Meeting of the ADI (October, 1964) pp. 421-423.

William Goffman and George Badger, Jr., *Parameters of Information Science*, "An Experiment with File Ordering for Information Retrieval", 1, from the *Proceedings* of the Annual Meeting of the ADI (October, 1964) pp. 379-381.

William Goffman and Vaun A. Newill, *Information Storage and Retrieval*, "Methodology for Test and Evaluation of Information Retrieval Systems", 3, pp. 19-25 (1964: AFOSR 65-0587; AD-614005).

William Goffman, *Nature*, "An Epidemic Process in an Open Population", 205, Number 4973, pp. 831-832 (February, 1965).

22

William Goffman, *Information Storage and Retrieval*, "On the Logic of Information Retrieval", **2**, pp. 217-220 (1965; AFOSR 65-2690, AD629801).

William Goffman, *J. Theoret. Biol.*, "A Mathematical Model for Describing the Compatibility of Infectious Diseases", **11**, pp. 349-361 (1966; AFOSR 67-1466, AD654236).

William Goffman, *Nature*, "Stability of Epidemic Processes", **210**, Number 5038, pp. 786-787 (May, 1966; AFOSR 67-1467, AD-654241).

William Goffman, *Nature*, "Mathematical Approach to the Spread of Scientific Ideas — The History of Mast Cell Research", **212**, number 5061, pp. 449-452 (October, 1966).

William Goffman and Vaun A. Newill, *Proc. Roy. Soc. Ser. A. London*, "Communication and Epidemic Processes", **298**, pp. 316-334 (1967; AFOSR 67-1474, AD654239).

Section 2

INFORMATION IDENTIFICATION AND CLASSIFICATION

The problem of selecting appropriate descriptors of information and organizing them into a structure was originally thought of as a library problem. It is still a crucial problem of information service activities handling the technical document literature. Conventional library classification schemes are not readily amenable to frequent and major modifications of narrow subject areas which must be accessed in specificity and depth as research expands knowledge of them. Conventional library indexing practices can similarly not readily accommodate to current requirements for many terminological access points to the subject content of documents.

The problem of identification and classification has, however, been aggravated by a technology that has produced computers and sensors. Machines that can read printed and handwritten characters and can look at maps and photographs need instructions to tell them how to recognize a pattern and how to differentiate one pattern from another. This raises the question of how to describe a pattern. What are the information-bearing parameters in a line drawing, in a letter of the alphabet, in an aerial photograph? Are the bits that are significant to a human significant to a machine? Can rules or algorithms be developed than can give machines the intelligence to identify patterns they haven't seen before and to separate signals from noise?

Projects concerned with both categories of problems pertaining to classification are included in this task. Descriptor structures for document systems are frequently related to other aspects of total system organization and operation. As methods are perfected for mechanizing useful indexing procedures, new approaches to input, search and retrieval can be explored and incorporated in systems. Work with natural language may also lead to abstractions pertaining to syntactic and sematic structures in language. Rules that characterize the process must describe procedures by which machines can learn from experience and can adjust or adapt to new inputs. How patterns are coded also bears on information content and information loss in their transmission.

The possibility exists that models for characterizing patterns will also describe concepts represented by clusters of word descriptors.

2.1 INFORMATION PROCESSING PERSONNEL SURVEY AND DATA BASE ANALYSIS

H.C. Asmus, Isaac D. Nehama

American Federation Of Information Processing
Societies (AFIPS), New York, New York

Current Contract Number* F44620-67-C-0092

Project Description

Increasing use and need by the federal government of complex
computer based systems have emphasized both the shortage of
competent personnel for these systems and the lack of comprehen-
sive documentation on available manpower from which plans and
projections for training, education, and utilization can be made.
This is a two-phase effort to acquire and analyze elements of the
needed data base. Personnel data will be obtained by questionnaire
from members of the Association for Computing Machinery, the
American Documentation Institute, the Association for Machine
Translation and Computational Linguistics, the Data Processing
Management Association, the Institute of Electrical and Electronics
Engineers Computer Group, the Simulation Councils, Inc., and
other groups found to be appropriate. Elements of the data base
will be examined with respect to their impact on the information
processing field (manpower, hardware, and software requirements;
education and training requirements; and government and industrial
application needs). Potential users and uses of the data base will be
identified, and recommendations will be made for the continued
development and augmentation of the base to meet current and
future user needs.

Summary of Recent Research Results

A Questionnaire has been distributed after a long delay in gaining
approval from the Bureau of the Budget. Answers to the survey
are now being received and tabulated.

26

2.2 OPTIMIZING THE STRUCTURE AND PERFORMANCE DESCRIPTIONS OF LEARNING NETWORKS

Johannes G. Goerner

Bell Aerosystems Co., Buffalo, New York

Contract Number AF-49 (638) 1449, AF-49 (638) 1627

Project Description

Most dynamic systems, physical systems, communication and other information processing systems, are highly nonlinear and are sensitive to a large number of variables. Recent research on multiple input learning systems suggests an approach to realizing the non-linearities through piecewise processing of inputs to the system, This effort continues a study of problems of discriminating nonlinearly separable patterns through the use of linear nets and nonlinear pre-processors. The use of binary logical operations and selected logic levels to improve performance of such nets having binary inputs was investigated. Necessary conditions were established to guarantee uniqueness of the transformed patterns for analog inputs. Experimental testing with a laboratory learning machine and a variety of preprocessors established the practical usefulness of different types of input coding. Both linear and nonlinear cascaded nets were also studied as an alternate method of improving discriminating properties of these networks. Due to the inherent limitations of fixed network structures, the effort is now directed at networks which have optimal structures based only on the probability densities of the pattern classes.

Summary of Recent Research Results

A simple solution to the important practical problem of the analysis of linear mixtures of constituent patterns has been found in the form of cascaded linear nets. It has been shown that the percentages of these patterns can be easily extracted from the mixed pattern.

All the network arrangements studied exhibit a prearranged structure. Although they have much practical merit, they nevertheless have limitations as the optimum decision boundary cannot always be obtained. Therefore, future research will be directed to an approach which leaves the net structure initially undefined and first determines the optimal performance based on known or unknown probability density functions of the patterns. The required structure of the net then becomes defined automatically. The problem areas

27

presently under study include developing the optimum structure network when: 1) The density functions of the input patterns are Gaussian with known variances and means; 2) The density functions of the input patterns are Gaussian, but the variances and means are not known and must be learned with the additional constraint of minimizing learning time; 3) The density functions of the input patterns are non-Gaussian, but known; 4) The probability density functions of the input patterns are non-Gaussian, but of known genral shapes defined by parameters which are not known and must be learned; and 5) The probability density functions of the inputs patterns have completely unknown shapes.

Reports and Publications

Johannes G. Goerner and Lester A. Gerhardt, "Analysis of Training Algorithms for a class of Self-Organizing Systems", in *Proceedings of the Symposium on Discrete Adaptive Processes*, Third National Electronics Conference, Chicago, Ill., 20-21 October 1964, pp. 631-636 (AFOSR 64-2538, AD 453-790).

Johannes G. Goerner, Lester A. Gerhardt and F. D. Powell, "A Unifying Mathematical Theory for Training Learning Nets, Final Report, October 1965" (AFOSR 65-2710, AD 627-908).

Johannes G. Goerner, Lester A. Gerhardt and F. D. Powell, "The Application of Error Correcting Learning Machines to Linear Dynamic Systems", in *Proceedings of the National Electronics Conference*, 1965, 21, pp. 541-546 (AFOSR 66-0012, AD 632-184).

Johannes G. Goerner, Lester A. Gerhardt and F. D. Powell, "Nonlinear Preprocessing of Inputs to Linear Neural Nets" (BAC Report No. 9500-920059, October, 1966; AFOSR 67-0054, AD 645-499).

Johannes G. Goerner, Lester A. Gerhardt and F. D. Powell, "Influence of Structure on Performance of Decision Networks" (BAC Report No. 9500-920089, October, 1967).

Lester A. Gerhardt and K.W. Drake, "Iteratively Generating the Convex Hull of Pattern Classes", submitted to *IEEE Transactions on Systems Science and Cybernetics*, September, 1967.

Lester A. Gerhardt, "A Learning System Using Multiaperture Cores" presented at *IEEE International Electronics Conference*, Toronto, Canada, September, 1967.

F.D. Powell, "A Nonlinear Adaptive Identifier", in *Record of the Southwestern IEEE Conference*, April 1967.

F.D. Powell, "A Closed Loop Parallel Structured Adaptive Modeling Net", submitted to *IEEE Group on Automatic Controls*, June, 1967.

F.D. Powell, "On Adaptive Nets" (Information Sciences Department Seminar, Univ. of Texas, April, 1965; BAC Report No. 9500-920083).

28

2.3 ANALYSIS OF PERSONAL INDEX STRUCTURES AND USES

Gerald Jahoda

Florida State University, Tallahassee, Florida

Project Description

Studies of the information gathering habits of scientists and engineers show that at least 50 percent maintain personal indexes which are often used more than library sources for information retrieval. This effort is a case history analysis of the personal index practices of a selected group of scientists to both characterize these practices and attempt to determine operating patterns by which documentation and information specialists and information centers could more usefully satisfy information needs. The effort includes study of the interface between the information specialist and the scientist, analysis of the structure and use of the index, and analysis of proposed modifications of various indexes and search inquiries. Results are expected to suggest methods of improving information gathering and retrieval for the individual scientist which should also reduce the time he uses in such operations.

Summary of Recent Research Results

Over 300 case histories of uses of personal indexes have been collected. The case histories consist of the questions and the bibliographic citations of relevant documents selected by the 12 researchers who participated in the study. Relevant documents selected in the course of the 300 index uses have been characterized in terms of type and age of documents and types of access points (author, subject, journal title) used. Keyword from title indexes have been prepared for the collections of documents in the offices of five researchers and case histories of uses of these five indexes are now being collected. Questions and relevant documents to these questions are also used in tests of keyword from title indexes for recall and precision ratios.

Reports and Publications

G. Jahoda, R.D. Hutchins, and D.M. Miller "Analysis of Case Histories Of personal Index Use", in *Proceedings of the 1966 Meeting of the American Documentation Institute*, 1966, pp. 245-54 (AD643869).

G. Jahoda, J.J. Oliva, and A.J. Dean, "Recall With Keyword From Title Indexes: Effect of Question-Relevant Document Title Concept Correspondence" (Library School, Florida State University, December, 1966).

G. Jahoda, R.D. Hutchins and R.R. Galford, *Am. Doc.*, "Characteristics and Use of Personal Indexes Maintained by Scientists and Engineers in One University", 17, No. 2, pp. 71-75 (April, 1966, AD644952).

30

2.4   LINEAR NETWORK AND PATTERN RECOGNITION

A. Timothy Ewald and Gordon Groves

University of Hawaii, Honolulu, Hawaii

Project Description

A primary tool used successfully in the study of artificial intelligence, in terms of pattern recognition, is the separation of data sets into discrete classes by the techniques of linear separability. These linear techniques, when applied to the multidimensional data space of a non-linear physical world, often lead to the development of data categories that are difficult to interpret in the physical world.

Although this effort proposes to employ both linear and quasi-linear techniques, the goal is to examine the identification process with respect to discovering differentiating descriptors and associations among the descriptors which can provide analytic relationships for describing the phenomenological world.

Point radar waveforms and Tsunami wave marograms are being used as the principal source of data. Preliminary results suggest that a relaxation of convergence criteria will lead to quantitative classification techniques for identifying detectible pattern information, possibly through the use of secondary signal cues. Results should be applicable to the construction of automatic decision systems and learning networks able to establish useful detectible signal categories in noisy systems.

Summary of Recent Research Results

This project has developed a family of algorithms appropriate to pattern recognition in terms of linear and quasi-linear separable sets and to clustering techniques by which a learning device can develop its own categories of data input in terms of realizability rather than *a priori* bounded categories.

Tsunami wave records were used as input data. The results of the tests show that the algorithms ares capable of detecting different classes of Tsunami waves with respect to their maximum energy characteristic. Heretofore, analytic mathematical techniques have been unsuccessful in this area.

Work has been carried out in developing "nearness" criteria on some 15 wave characteristics to determine the effect of each parameter on the "detectability" of clusters. Along with this aspect of the work, the learning system has been extended to produce its own detectable clusters without the benefit of learning trials with known prototypes. We are now trying to relate the clustering behavior of the learning system to the physical parameters of the system.

Of mathematical interest is the demonstration that a general error correcting learning algorithm has as a special case a learning scheme developed by Professor M. Watanabe.

Reports and Publications

A. Timothy Ewald, "Classification and Prediction of Tsunami Waves by Clustering and Threshold Techniques", presented at the *IEEE International Symposium on Information Theory*, San Remo, Italy, 1967.

A. Timothy Ewald, "Pattern Recognition in Classification of Seismic Disturbances", to be prepared and submitted to the Geological Society of America.

A. Timothy Ewald, "Pattern Recognition and Monopulse Radar", to be prepared and presented to the Jour. of IEEE.

A. Timothy Ewald, "Wave-Parameter Clusters", to be prepared and presented to the Jour. of Geographic Research.

2.5 SELECTION OF VARIABLES IN CLUSTERING AND PATTERN RECOGNITION

Michael S. Watanabe

University of Hawaii, Honolulu, Hawaii,

Project Description

Part of the problem involved in specifying procedures for the machine recognition of patterns derives from the existence of n possible independent predicates to describe the results of observations and the existence of 2^{2^n} possible predicates derivable from them. The prot'em resides in the selection of the few predicates from the possible ones that are important for classification. This effort will examine bases for improving current approaches to automatic pattern recognition and for establishing new measures of "goodness" of variables, particularly for the pre-selection of variables toward solving clustering and recognition problems of classification. Incorporation of the investigator's computer-feasible method of pre-weighting, called SELFIC, in a clustering algorithm will be evaluated. Methods are sought for extracting basic invariant features of characters that identify simple variables approximating those used by human recognizers. A potential value assignment for points in n dimensional space will be examined that yields unique, reasonable classification solutions. A convergence method is being studied that tends to select a hyperplane in a well-balanced position between two groups of points.

Summary of Recent Research Results

It is universally recognized that the secret of successful algorithms of clustering and pattern recognition depends on a good choice of variables in terms of which objects are represented. Generally speaking, there is no unique criterion by which the "goodness" of variables is to be evaluated. However, there is one way of deriving a unique criterion of evaluation solely from the statistical properties of the ensemble of objects at hand. This is the method of SELFIC which has proved to be very powerful in many pattern recognition problems (1,2,3).

Under the present contract, this method was further developed and a new computer program was developed. This program has, among other advantages, the capability of handling vectors with many more components than was previously possible. The SELFIC method is found to be useful as the first step in information compression even in those cases where the final criteria for good variables do not originate from the ensemble itself.

33

The most significant achievement of the project during the first contract year was the application of the SELFIC method to the clustering algorithm. We had anticipated that the SELFIC method would be instrumental in reducing the number of variables to a small number of important variables. However, the use of SELFIC for the purpose of reducing the number of objects was an unexpected innovation that happened during this sponsorship. In fact, most of the clustering algorithms that have been proposed suffer from a common shortcoming (4). Namely, as soon as the number of objects grows, say over 20, the computation takes a prohibitively long time. This common difficulty can now be overcome by the special use of SELFIC. Essentially, we select a small number of important variables, and those objects that have large values in these variables are selected as representative objects to which we may apply the clustering algorithm to generate classes (5). After the classes are established we may reinstate non-representative objects and place them in these classes by the use of a "recognition" algorithm.

This method is so promising that we intend to develop and improve it in the second contract year.

Reports and Publications

1. S. Watanabe, "Karhumen-Loeve Expansion and Factor Analysis–Theoretical Remarks and Application–" in the *Proceedings of the Fourth Prague Conference on Information Theory, 1965* (Czechoslovakia Academy of Science, Prague, 1966).
2. S. Watanabe, "A Method of Self-Featuring Information Compression in Pattern Recognition," in the *Proceedings of the BIONICS Conference 1966* (in press).
3. S. Watanabe, *et alia*, "Evaluation and Selection of Variable in Pattern Recognition" in the *Proceedings of the COINS Conference, 1966* (Academic Press, 1967).
4. S. Watanabe, "Mathematical Explication of Classification of Objects," in *Information and Prediction in Science*, Dockx, Bernays, eds., (Academic Press, 1965).
5. S. Watanabe, *et alia*, "Clustering Algorithm with Automatic Preselection of Variables," presented at the IEEE San Remo Conference on Information Theory, 1967.

34

2.6 UNIFICATION OF THEORY AND EMPIRICISM IN INFORMATION RETRIEVAL

Robert A. Fairthorne

Herner and Company, Washington, D.C.

Contract Numbers AF-49 (638) 1261,
AF-49 (638) 1427, AF-49 (638) 1617

Project Description

Numerous theoretical and empirical efforts have been made to descriptively characterize and organize informational data for subsequent retrieval. Many of these efforts contain unclear and/or erroneous hypotheses concerning the factors involved in retrieval, dissemination, and information networks in general. The goal of this effort is to identify the useful targets for research in Information Retrieval (IR) by establishing the limits of the field. By identifying the useful targets for observation and experimentation in the IR field, one helps eliminate work on tasks that cannot possibly be performed, or even if possible, would be worthless. The limits of IR are fixed, however industrious the researchers, and however ingenious their methods and devices. Therefore, an important result of their identification is the establishment of suitable criteria of evaluation that are both possible and within the scope of IR systems. The general research plan is to develop further the operational approach devised in previous research efforts. This approach identifies the physical objects involved, the circumstances of their production, and the circumstances of their use. The objects of interest are texts whose production involves choices of marks and media on which to record them. It is hoped that this analysis and description of IR will help identify the operations that are inevitable and those that allow choice.

Summary of Recent Research Results

The scope of retrieval is that of management of previously recorded discourse for the benefit of other readers. Thus it falls short of active participation in discourse on one hand, and beyond mere signalling or recording on the other.

Under this study a clarification of terminology has been achieved by pointing out that Information Systems neither originate nor use the records they manage, but only mention them to readers. The simplest form of Information System is a Notification System (i.e. one that supplies or mentions to readers such recorded messages that may satisfy their specifications). These involve Message, Source, Destination, Designation, Code and Channel.

It cannot consider less than three of these at a time, and the twenty resulting triads are the basic activities of any system. "Flow" between two of the six variables is possible only when the two are members of distinct triads with two variables in common. Then flow between any two variables of this tetrad entails a corresponding flow between the conjugate pair. With properly defined variables, Shannon entropic measures are valid, but do not measure the same things. For instance, there are six distinct pairs of variables that may form a tetrad with "Destination" and "Message". Each is associated with a distinct type of "Information Flow".

This research effort has also shown the some current numerical measures of retrieval performance are highly dependent on the composition of the particular collection tested, and therefore are useless for comparison or prediction.

Reports and Publications

Robert A. Fairthorne, "Mathematics, Mechanics and Statistics for the Information Science Curriculum or What Mathematics Does an Information Scientist Need", in *Proceedings of the 26th Annual Meeting of the ADI*, Chicago, Illinois, pp. 39-40 (October 1963; AFOSR 64-0238, AD432794).

Robert A. Fairthorne, *J. Doc.*, "Review of Self-Organizing Systems 1962", 19, Number 4, pp. 204-205 (December 1963; AFOSR 64-0232, AD432520).

Robert A. Fairthorne, "Basic Parameters of Retrieval Tests", in *Proceedings of the ADI*, 1, Philadelphia, pp. 343-345 (October 1964; AFOSR 65-0593, AD616175).

Robert A. Fairthorne, *J. Doc.*, "Review of Natural Language and the Computer 1963", 21, Number 3, pp. 201-205 (Septamger 1965).

Robert A. Fairthorne, "The Necessity of Non-Descript Descriptors", in *Proceedings of the ADI*, 1, Philadelphia, pp. 355-356 (October 1964).

Robert A. Fairthorne, "Basic Parameters of Retrieval Tests", in *Proceedings of the ADI*, 1, Philadelphia, pp. 343-345 (October 1964; AFOSR 65-0593, AD616175).

Robert A. Fairthorne, "Use and Mention in the Information Sciences", in *Education for Information Science* (Spartan Books, Washington, D.C., 1965; AFOSR 65-2289, AD629367) pp. 9-12.

Robert A. Fairthorne, *J. Doc.*, "Some Basic Comments on Retrieval Testing", 21, number 4, pp. 267-270 (December 1965; AFOSR 66-0414, AD632181).

Robert A. Fairthorne, *J. Ass. Comp. Soc.*, "Morphology of Information Flow", 14, number 4, pp. 710-719 (October 1967).

Robert A. Fairthorne, "The Applied Mathematics of H.P. Luhn" in *Hans Peter Luhn Pioneer and Prophet of Information Processing*, New York, N.Y. (October 1967).

Robert A. Fairthorne, "Information Processing Historical Sketch", in *Encyclopedia Britannica*, Chicago (1968).

2.7 ANALYSIS OF DOD SPONSORED RESEARCH PUBLICATIONS

Irving H. Sher

Institute for Scientific Information
Philadelphia, Pennsylvania

Project Description

This research program will capitalize upon the data store which has been compiled by the Institute for Scientific Information, regarding citations of articles in the scientific literature. In 1964, all articles entered into the ISI data store were coded according to the agency or organization which sponsored the research leading to the article. All articles sponsored by the Air Force Office of Scientific Research, the Army Research Office, and the Office of Naval Research and published in 1964, will be checked for subsequent citations in the scientific literature. If a subsequent article makes reference to one of the sponsored articles, it can be assumed that the sponsored article was of some value to the author of the citing article. Patterns of citation for different types of articles, for different journals, etc., will be examined to see what conclusions can be drawn which might lead to improved management of research programs. Use of articles sponsored by all three Service agencies will give a larger data base than would use of one agency's articles alone.

Summary of Recent Research Results

The open literature and literature of technical reports were reviewed and the available methods for producing and displaying graphic networks were evaluated. In the light of the problem of handling large networks (up to 100 nodes or more) with many interconnections (unlimited connections among nodes). a new technique was developed called "Diagonal Display" and which appears to be very promising as a method of mechanically producing networks by printout, by computer directed plotters or for electronic display tubes. Diagonal Displays can be manipulated by relatively simple algorithms and have the property of being indefinitely extensible. There is no limitation in Diagonal Display on the number of possible interconnections among any of the nodes, and Diagonal Display retains a simple clarity even when the number of interconnections grows large.

2.8 A FORMAL THEORY OF CONCEPTUAL AFFILIATION FOR DOCUMENT RECONSTRUCTION

Donald J. Hillman

Lehigh University, Bethlehem, Pennsylvania

Project Description

Grammatical analysis has potential value for problems of text processing, i.e., document retrieval and message retrieval. However, a grammar suitable for machine manipulation must have information preserving, meaning preserving, and truth preserving properties. The grammar should also be notationally and algorithmically simple. Previous research indicates the necessity of combining logical and epistemological investigations with linguistic study to develop the fundamental concepts for a grammar for retrieval purposes. This effort is for the formulation of a theory of conceptual affiliation to supply the logical foundation of a process by which characteristic terms are organized into coherent document-forming sets. Logical refinement of the notions of term-term and document-term association should aid in the construction of a grammar suitable for machine text processing.

Summary of Recent Research Results

Research has centered upon the significant logical and epistemological issues involved in the development of an adequate text-processing procedure for document retrieval systems, and it was shown that the problem of characterizing or indexing a document is a particular form of the general logical problem of abstraction. The notion of *concretion* has been introduced into retrieval theory, and the problem of reconstructing documents from their given characteristics has been examined. A major goal was to show that different coherence patterns for characteristics indicate differences among documents as to form, style, structure and content.

An affiliation-value model was constructed, in terms of which the conceptual closeness and importance of characteristics to their parent documents are numerically expressible. Document structure, organization and content are reflected in the calculation of affiliation-values, so that it becomes possible to identify the form and nature of documents given only certain minimal information regarding their characteristics or index terms.

38

Reports and Publications

Robert C. Heiser and Donald J. Hillman, "The Affiliation-Value
  Model, A Formal Theory of Conceptual Affiliation for Docu-
  ment Reconstruction" (Report No. 1, August 22, 1966).
Robert C. Heiser and Donald J. Hillman, "Reconstruction Models"
  (Report No. 2, in preparation).

39

2.9 **TOPICS IN PATTERN RECOGNITION
AND COMMUNICATION SCIENCE**

Laveen N. Kanal

Lehigh University, Bethlehem, Pennsylvania

Project Description

Fundamental information processing problems are disclosed in
present attempts to perfect machine-based systems for pattern rec-
ognition, speech processing, visual communication, tactical com-
munication, and high-speed data transmission. Problem solution re-
quires a combination of advanced mathematical tools and the formu-
lation of new concepsual approaches to information classification,
compression, and coding. This effort will examine and refine the-
oretical constructs needed for improved information processing sys-
tems. Topics will include analysis of sample size in statistical pat-
tern recognition methods, measures for the complexity of patterns,
classification functions and decision procedures for patterns whose
geometric features are not obvious, infinite-state Markov models for
communication channels with memory, classification of nonstation-
ary time series, and information diffusion in large distributed com-
munication networks. Measures, models and feasibility tests are ex-
pected to advance real system capabilities in pattern discrimination
and functioning of impaired communication networks.

Summary of Recent Research Results

This AFOSR research effort received its initial sponsorship less
than one year ago, and will be reported in the next issue of this pub-
lication.

AIR FORCE SCIENTIFIC RESEARCH BIBLIOGRAPHY

G. Vernon Hooker and Clement R. Brown

Library of Congress, Washington, D.C.

Project Description

Needs of research scientists for comprehensive, authoritative bibliographic tools and needs of administrative scientists for complete, compact descriptions of the programs they manage are well known. This effort is to satisfy both needs with respect to Air Force-sponsored (in-house and contractual) basic research. This effort provides for the compilation of abstracts and other descriptive information pertaining to the reports issued under the Air Force basic research program. The compilations are being published in volumes with the entries arranged by contractor, and with subject contract, personal author, control number indexes, and a special index to mathematical publications. Presently these are appearing on a yearly basis.

Summary of Recent Research Results

Since January 1966, Volume IV, covering 1960, 909 pages with 3047 items has been published and Volume V, covering 1961, which has 3207 items and will include a cumulative author index for the first five volumes will be submitted to the printer in October 1967. In addition, many items for future volumes have been collected and abstracted.

Reports and Publications

G. V. Hooker, M. H. Duffner, A. S. Dann, and D. C. Yates, *Air Force Scientific Research Bibliography 1950 - 1956* (U. S. Government Printing Office, Washington, D.C., 1961). **1**, AFOSR 700.

G. V. Hooker, A. S. Dann, D. C. Yates and M. S. Carr, *Air Force Scientific Research Bibliography 1957 -1958*, (U. S. Government Printing Office, Washington, D.C., 1964), **2**, AFOSR 700-II, AFOSR 64-0967.

G. V. Hooker, D. C. Yates, H. D. Brookins and A. S. Dann, *Air Force Scientific Research Bibliography 1959*, (U. S. Government Printing Office, Washington, D.C., 1965), **3**, AFOSR 700-III.

G. V. Hooker, D. C. Yates, H. D. Brookins, J. E. Halpin and P. A. Patrick, *Air Force Scientific Research Bibliography 1960,* (U. S. Government Printing Office, Washington, D.C., 1966), **4**, AFOSR 700-IV.

G. V. Hooker, D. C. Yates, H. D. Brookins, J. E. Halpin, P. A. Patrick and J. D. Weber, *Air Force Scientific Research Bibliography 1961,* (U. S. Government Printing Office, Washington, D.C., In Press), **5**, AFOSR 700-V.

42

2.11   THEORETICAL FRAMEWORKS FOR
PATTERN RECOGNITION PROBLEMS

Wilson P. Tanner Jr.

University of Michigan, Ann Arbor, Michigan

Project Description

Much research on pattern recognition has been pragmatic attempts to devise suitable techniques for recognizing elements of a particular set of pattern ensembles. In this effort to develop theoretical frameworks, the aim was to develop the following operations that are implicitly assumed to be performed by man as a pattern recognizer: (1) input weighting and storage, (2) decomposition on a mixed distribution of input storage; (3) estimation of the parameters of the distributions determined by the decomposer (the parameters defining the elementary codes from which patterns are constructed); (4) statement of the pattern set and associated probabilities based on all (distant and recent) prior states of knowledge; and (5) storage of experience.

Summary of Research Results

This project has been directed toward basic studies of models for pattern recognition. Initially the effort emphasized general models but more recently specific problems within the framework of the general model have been attacked. One such problem has resulted in a study of the decomposition of summed distributions; another problem studied within the framework of the general model is that of experimental design recognizing that the interpretation of scientific data involves the same processes as the interpretation or recognition of patterns.

Reports and Publications

R.F. Arnold and D.L. Richards, "Monotone Reduction Algorithms", in *Proceedings of the International Conference on Microwaves, Circuit Theory, and Information Theory*, Tokyo Japan, September 1964, pp. 175-176 (AFOSR 65-0586; AD 619281).
R.F. Arnold and D L. Richards, "Monotone Congruence Algorithms" (Supplementary Technical Report Number ISL-65-2, April 1965; AFOSR 65-1736, AD625046).

J.B. Ranck, Jr. and S.L. Bement, *Experimental Neurology*, "The Specific Impedance of the Dorsal Columns of Cat: An Anisotropic Medium", 11, Number 4, pp. 451-463 (April 1965; AFOSR 65-1166, AD 620527).

D.L. Richards, "An Abstract Model for the Pattern Recognition Process (Technical Report Number 121, Cooley Electronics Laboratory, April 1961).

D.F. Stanat, "Nonsupervised Pattern Recognition Through the Decomposition of Probability Functions (Supplementary Technical Report, April 1966; AFOSR 66-1454, AD637486).

W.P. Tanner, Jr., "Pattern Recognition and Information Processing (Final Technical Report, January 1965; AFOSR-0242, AD 611135).

W.P. Tanner, "Criteria for Laboratory Experiments Useful in Field Situations", in *Vision Research - Flying and Space Travel*, National Academy of Sciences, Washington, D.C. (to be published).

2.12 ANALYSIS OF INFORMATION STORAGE, SEARCH, AND RETRIEVAL PROCESSES

Hans Selye

University of Montreal, Montreal, Canada

Project Description

A notation system for representing detailed scientific information content, called the Symbolic Shorthand System, was examined for its amenability to subject expansion and its utility for information retrieval. A unique data base in the medical science literature was the test corpus. Experiments were conducted to qualitatively evaluate the documentation service based on the notation system. The experiments were designed within the environment of on-going service operations.

Summary of Research Results

This project enabled the Institute of Experimental Medicine and Surgery to use its documentation services to prepare bibliographies on specialized subjects for scientists whose primary research interests were related to the fields of endocrinology, stress, mast cells, anaphylactoid inflammation, calciphylaxis and calcergy, cardiac necroses, thrombohemorrhagic lesions and calcitonin (thyrocalcitonin). Of the 199 scientists originally contacted, 114 showed interest in receiving the bibliographies, including 22 attached to our Institute.

A total of 25,152 references were furnished and on the basis of responses to prepared evaluation sheets, 86% of these were established as pertinent. Many users expressed surprise at the degree of precision attained in the bibliographies, even in the case of highly specialized subjects. For instance, in the field of calcitonin, such internationally recognized investigators as Drs. D.H. Copp and I. MacIntyre qualified it as the most complete and up-to-date reference collection available anywhere.

In addition, the project provided an opportunity to compile extensive indexes on the subjects of prime interest at our Institute. The reference material thus obtained proved highly valuable in the preparation of the monographs: *THROMBOHEMORRHAGIC PHENOMENA* (Springfield: Charles C. Thomas Publ., 1966) and *ANAPHYLACTOID EDEMA* (St. Louis: Warren H. Green Inc., in press).

Reports and Publications

George Ember and G. Gabbiani, *Bull. A.I.D.* "The Documentation Service of the Institute of Experimental Medicine and Surgery of the University of Montreal", 2, Number 4, pp. 19-22 (1963; AFOSR 64-0683, AD 436517).

N. Padmanabhan and George Ember, *Meth. Info. Med.*, "Symbolic Shorthand System for Physiology and Medicine, 1, Number 4, pp. 138-142 (October 1962; AFOSR J 711; AD415397).

Hans Selye and George Ember, *Symbolic Shorthand System (SSS) for Physiology and Medicine* (Institute of Experimental Medicine and Surgery, University of Montreal, 1964; AFOSR 64-0681).

Hans Selye, "Information Retrieval Experiment" (Final Technical Report, June 1967; AFOSR 67-2362, AD 659716).

2.13 DIGITAL COMPUTER PROCESSING
OF GRAPHICAL DATA

Herbert Freeman

New York University, New York, New York

Project Description

The ability to use digital computers for the analysis, synthesis, and manipulation of graphic contour-line data can have applicability in the analysis of complex geometric patterns as are found on maps, optical and radar signal patterns, on other printed and handwritten images. Research under this project has developed efficient techniques of encoding geometric information and algorithms for the recognition, analysis, and classification of geometric patterns. During the past year, a technique was developed for solving the generalized two-dimensional allocation problem. A problem in this category would include finding the optimum manner in which irregular shapes may be cut from fixed-sized sheets. Another would be the "parking-lot problem" in which it is desired to find the maximum number of automobiles that can be parked in an irregularly shaped lot. One area of investigation will be to study the effect on these techniques when additional parameters are introduced in order to substantiate the belief that these techniques may be exstensible. Another area of investigation is the application of linguistic analysis to computer graphics. A set of productions will be developed in an attempt to direct pattern analysis using linguistic structures.

Summary of Recent Research Results

A method has been developed for solving the optimum two-dimensional allocation problem. The method consists of two parts. In the first, various combinations of given irregular shapes are grouped into clusters and circumscribed with rectangles in an optimum manner. In the second, the simplex method of linear programming is employed to find the optimum combination and arrangement of these rectangles that satisfy the constraints of the problem. Speed of convergence is increased by means of a dynamic programming algorithm which generates new columns for the simplex tableau at each iteration.

An investigation has been carried out to explore the possibility of finding a linguistic structure for graphical patterns. It appears that

47

much of the theory that has been developed for conventional one-dimensional languages (both natural and artificial) can be extended to pattern languages. It was possible to show that linguistic techniques can be applied to the specification of pattern classes and to the structural analysis of patterns. It was shown that graphical patterns possess a definite language syntax. The specifications of a parameterized compiler have been developed. When completed, this compiler, called a parameterized pattern analyzer, will be able to perform a variety of pattern analysis tasks on widely differing classes of patterns.

Reports and Publications

P.T. Brady, "Queueing and Interference Among Messages in a Communication System with Transmission Delay" *(Doctoral dissertation,* Department of Electrical Engineering, New York University, NYU Tech Report 400-134).

Jerome Feder and Herbert Freeman "Segment Fitting of Curves in Pattern Analysis Using Chain Correlation" (Technical Report No. 400-108, March 1965; AFOSR 65-1206, AD619525).

Jerome Feder, "A Linguistic Approach to Pattern Analysis: A Literature Survey" (Technical Report No 400-133, February 1966; AFOSR 66-1802, AD637497).

Jerome Feder and Herbert Freeman "Digital Curve Matching Using A Contour Correlation Algorithm" in *IEEE Intern. Conv. Record,* pp. 69-85 (1966; AFOSR 66-2185, AD644577).

Jerome Feder, "Linguistic Specification and Analysis of Classes of Patterns" (Technical Report 400-147, October 1966; AFOSR 67-0987, AD651606).

Herbert Freeman, "On the Digital Computer Classification of Geometric Line Patterns," in *Proc. Nat'l. Elect. Conf.,* 18, pp. 312-324, Chicago, Ill. (October 1962).

Herbert Freeman, "Pattern Analysis of Planar Geometric Configurations" (Progress Report, October 1962- November 1966).

Herbert Freeman and L. Garder *IEEE Trans Electron Computers,* "Apictorial Jigsaw Puzzles - The Computer Solution of a Problem in Pattern Recognition", 13, Number 2, pp. 118-127 (April 64; AFOSR 64-1386, AD444464).

Herbert Freeman "Pattern Analysis of Planar Geometric Configurations" (Final Report, December 1964).

Herbert Freeman, "Pattern Analysis of Planar Geometric Configurations" (Final Report, December 1965).

Herbert Freeman, "Pattern Analysis of Planar Geometric Configurations" (Final Report, April 1966; AFOSR 66-0776, AD632365).

Herbert Freeman, "On the Classification of Line-Drawing Data," in *Models for the Perception of Speech and Visual Form,* Wathen-Dunn ed. (M.I.T. Press, Cambridge, Mass. 1967).

Herbert Freeman and Jerome M. Glass, "On the Quantization of Line-Drawing Data," submitted for publication to the *Journal of the Association for Computing Machinery.*

48

Herbert Freeman and P. Loutrel, *IEEE Trans. Electron. Comp.*, "An Algorithm for the Solution of the Two-Dimensional 'Hidden-Line' Problem", **EC-16**, number 6 (December 1967).

Jerome M. Glass, "The Quantization of Two-Dimensional Line Drawings" (Technical Report 400-66, Department of Electrical Engineering, New York University, February 1963).

Jerome M. Glass, "A Criterion for Quantization of Planar Contours" (Technical Report, NYU TR 400-91, February 64; AFOSR 64-0940, AD600310).

Jerome M. Glass, "A Criterion for the Quantization of Line-Drawing Data" (Technical Report Number 400-112, May 1965; AFOSR 65-1495, AD621086).

Jerome M. Glass, *NORD. TIDSKR, Inform. Behandling*, "Smooth-Curve Interpolation a Generalized Spline-Fit Procedure" **6**, Number 4, pp. 277-293 (1966; AFOSR 67-0728, AD652437).

M.J. Haims, "On the Optimum Two-Dimensional Allocation Problem" (Technical Report 400-136, June 1966; AFOSR 66-2698, AD642761).

P. Kaszerman, "Properties of Threshold Functions Derived by the Use of Arthur's Geometric Map" (Technical Report 400-63, Department of Electrical Engineering, New York University, July 1962).

P. Kaszerman, "Synthesis of a Nonlinear Summation Threshold Device" (Technical Report 400-64, Department of Electrical Engineering, New York University, July 1962).

P. Kaszerman, "A Geometric Approach to the Study of Threshold Functions" (Technical Report 400-70, Department of Electrical Engineering, New York University, January 1963).

P. Kaszerman, "On the Synthesis of Threshold Devices" (Technical Report 400-84, September 1963; AFOSR 64-0394.

P. Kaszerman, *Inform. Control*, "A Geometric Test-Synthesis Procedure for a Threshold Device", **6**, Number 4, pp. 381-398 (December 1963; AFOSR 64-0391, AD435652).

P. Kaszerman, *IEEE Trans. Electron. Computers*, "A Nonlinear Summation Threshold Device", **EC-12**, pp. 914-915 (December 1963; AFOSR 64-1391, AD444462).

P. Kaszerman, *Inform. Control*, "A Region Concept and Its Application to Threshold Logic", **8**, Number 5, pp. 531-551 (October 1965, AFOSR 65-2684, AD628827).

S.P. Morse, "Computer Storage and Analysis of Contour Map Data" (Technical Report Number 400-106, February 1965; AFOSR 65-0592, AD617955).

S. Morse, "Analysis of a Contour Map on a Closed Surface" (Technical Report 400-123, Department of Electrical Engineering, New York University, September 1965).

S. Morse, "A Mathematical Model for the Analysis of Contour-Line Data" (Technical Report 400-124, Department of Electrical Engineering, New York University, October 1965).

M. Raship, "m-out-of-n Decision Logic" (Technical Report 400-85, Department of Electrical Engineering, New York University, October 1963). (Also published in *Proc. National Electronics Conference*, **19**, pp. 131-141, October 1963).

M. Raship, "On the Analysis and Synthesis of Switching Networks Composed of m-out-of-n Decision Gates", (Technical Report Number 400-95, June 1964; AFOSR 64-1378, AD604043).

M. Raship, "On the Realization of Binary Switching Functions With Majority Gates" (Technical Report Number 400-100, September 1964; AFOSR 64-2495, AD609504).

M. Raship, "An Algebraic - Geometric Technique for the Realization of Switching Functions With m-out-of-n Decision Gates" (Technical Report No. 400-101, October 1964; AFOSR 64-2496, AD609505).

K. Ruttenberg, "Digital Computer Analysis of Arbitrary Three-Dimensional Geometric Configurations" (Technical Report 400-69, Department of Electrical Engineering, New York University, October 1962).

K. Ruttenberg, "Algorithms for the Encoding of Three-Dimensional Geometric Figures" (Technical Report Number 400-86, June 1963; AFOSR 64-0560, AD434320).

Eliott B. Socci, "Program for the Geometric Analysis of Encoded Line Patterns" (Technical Report 400-78, Department of Electrical Engineering, New York University, April 1963).

Eliott B. Socci, "A Study of the Chain-Difference Correlation Technique for Contour Line Patterns" (Technical Report 400-22, March 1964; AFOSR 64-0941, AD600276).

50

SPEECH COMMUNICATION AND AUTOMATIC SPEECH RECOGNITION

Gordon E. Peterson* and June E. Shoup

Speech Communications Research Lab. Inc.,
Santa Barbara, California

Project Description

The realization of speech communication with an electronic device such as a computer depends on the computer ability to accept and interpret speech input (automatic speech recognition) and to produce meaningful speech output (automatic speech synthesis). Before the realization of these goals is possible, an investigation of the fundamental properties of spoken language, and a study of the principles basic to the use of electronic instrumentation in speech communication is necessary. This effort seeks to develop a basic understanding of the essential information bearing parameters of speech in order to transform the acoustic parameters of speech into a discrete code amenable to computer processing. Research on this effort has produced a general phonetic theory. Research is continuing on: a) a theory of phonology with an emphasis on phonemic and prosodemic theory and the relation between phonology and grammar, including lexicon and syntax; b) logical procedures for interpreting acoustic phonetic data such as wave types and acoustical parameters of speech; c) a structural description of the phonology of Midwestern American English which is necessary to convert phone types to orthographic sequences; and d) the development of a lexical theory in order to isolate and identify words within orthographic sequences.

Summary of Recent Research Results

An extensive amount of work has been conducted on phonemic theory and a formalization of this theory is near completion. Research on the intonation of French which employed both speech synthesis and acoustic speech analysis has contributed to our work on prosodemic theory. It is hoped that the final results of the study will indicate the relationship that fundamental voice frequency and acoustic speech power have to judgments of intonation. A computer program is being written for the logical procedures which were previously developed for the interpretation of acoustic phonetic data.

---

*Deceased.

Refinements have been made in the statement of the phonemes and allophones of Midwestern American English and in the tabulation of the phonetic environmental conditions under which the various allophones of the phonemes occur. Further work has also been done on the detailed tabulation of the allowable phoneme sequences of Midwestern American English. Six technical articles and monographs have been published since January of 1966.

Reports and Publications

J.A. Mason and G.E. Peterson, *Language and Speech,* "On the Problem of Describing the Grammar of Natural Language", **10,** Part 2, pp. 107-121 (April-June 1967).

M.H. O'Malley and G.E. Peterson, *Phonetica,* "An Experimental Method for Prosodic Analysis", **15,** No. 1, pp. 1-13 (1966).

G.E. Peterson, "Linguistic Concepts in Automatic Speech Recognition Procedures", in *Proceedings of the Seminar on Speech Compression and Processing,* **2,** AFCRC-TR-59-198 (1959).

G.E. Peterson and F. Harary, "Foundations of Phonemic Theory", in *Proceedings of Symposia in Applied Mathematics,* **12,** pp. 139-165 (1961).

G.E. Peterson and N.P. McKinney, *Phonetica,* "The Measurement of Speech Power", **7,** No. 2/3, pp. 65-84 (1961; AD611293).

G.E. Peterson, *Language and Speech,* "Automatic Speech Recognition Procedures", **4,** Part 4, pp. 200-219 (October-December 1961; AD453604).

G.E. Peterson and C.J. Fillmore, "The Theory of Phonemic Analysis", in *Proceedings of the Fourth International Congress of Phonetic Sciences,* Helsinki (Mouton and Co., 1962) pp. 476-489.

G.E. Peterson and J.E. Shoup, "Report on Attendance at European Conferences on Acoustics and Speech, 16 August- 14 September 1962" (Scientific Report, October 1962; AFOSR 3882, AD289314).

G.E. Peterson, "Current Problems in Speech Communication" (Communication Science Seminar Florida U, Gainesville, Florida, 12 June 1963; AFOSR 5068).

G.E. Peterson, "Research on Speech Analysis" (Final Report, October 1963; AFOSR 1162).

G.E. Peterson, *Phonetica,* "Requirements for a Phonetic Theory" **11,** No. 3-4, pp. 228-236 (1964).

G.E. Peterson, "Research on Speech Communication and Automatic Speech Recognition" (Final Report, January 1966; AFOSR 66-0186, AD478122).

G.E. Peterson and J.E. Shoup, *Journal of Speech and Hearing Research,* "A Physiological Theory of Phonetics", **9,** No. 1, pp. 5-67 (March 1966).

G.E. Peterson and J.E. Shoup, *Journal of Speech and Hearing Research,* "The Elements of an Acoustic Phonetic Theory", **9,** No. 1, pp. 68-99 (March 1966).

G.E. Peterson and J.E. Shoup, *Journal of Speech and Hearing Research,* "Glossary of Terms from the Physiological and Acoustic Phonetic Theories", **9**. No. 1, pp. 100-120 (March 1966).

G.E. Peterson and J.E. Shoup, "Research on Speech Communication and Automatic Speech Recognition" (Final Report, August 1966; AFOSR 66-2692, AD642763).

G.E. Peterson, "On the Nature of Speech Science", Annual Bulletin No. 1, Research Institute of Logopedics and Phoniatrics, University of Tokyo, pp. 13-19 (1967).

G.E. Peterson, "The Relation of Descriptive Phonetics to Research in Speech Communication" (accepted for publication in Glossa).

G.E. Peterson, "Toward a Theory of Symbolization" (to be published in a memorial volume for Elwood Murray).

G.E. Peterson, "The Speech Communication Process", *Manual of Phonetics,* B. Malmberg, Ed. (North-Holland Publishing Company, Amsterdam, Netherlands; revised edition) accepted for publication.

M. Rothenberg, *Bibliotheca Phonetica,* "The Breath-Stream Dynamics of Simple-Released-Plosive Production", Fasc. 6, (1967).

J.E. Shoup, *The Journal of the Acoustical Society of America,* "Phoneme Selection for Studies in Automatic Speech, Recognition", **34,** No. 4, pp. 397-403 (April 1962; AD453603).

W. S-Y. Wang and J. Crawford, *Language and Speech,* "Frequency Studies of English Consonants" **3,** Part 3, pp. 131-139 (July-September 1960).

W. S-Y. Wang and C. J. Fillmore, *Journal of Speech and Hearing Research,* "Intrinsic Cues and Consonant Perception", **4,** No. 2, pp. 130-136 (June 1961).

2.15 COMPENDIUM OF THE DISTRIBUTIONS OF MATHEMATICAL STATISTICS AND APPLICATIONS

Samuel Kotz

University of Toronto, Toronto, Canada

Project Description

Mathematical statistics is becoming increasingly important to the information sciences in furnishing tools for the representation of research results. This effort is for the compilation of statistical distributions which will serve as a bibliographic source and a handbook useful to those doing research in statistics and in applications of statistics. The compendium will contain compact but detailed descriptions of the distributions, properties of the estimated parameters, and examples from major areas of their applications. Source materials for the compendium include a substantial number of non-U.S. papers, including Russian and Polish materials, analyzed in the original languages. Over 100 S.K. distributions and their variations and generalizations are being considered.

Summary of Recent Research Results

Part I of the project which deals with discrete distributions was completed in March 1967. This work consists of 11 chapters dealing in a unified form with some 45 discrete distributions and their variations. Over 1500 papers have been consulted during the preparation and of those over 600 References are included. This volume is scheduled for publication in 1968.

Volume II dealing with a unified continuous and organized treatment of univariate distributions is in final stages of preparation. This volume also contains a long (75 typed pages) chapter on the normal distribution. Of the 1000 papers consulted only 250 were found suitable for inclusion in the bibliography.

Samuel Kotz, N.L. Johnson and D.W. Boyd, *Ann. Math. Statist.*, "Series Representations of Distributions of Quadratic Forms in Normal Variables I Central Case", **38**, (1967).

Samuel Kotz, N.L. Johnson and D.W. Boyd, *Ann. Math. Statist.*, "Series Representations of Distributions of Quadratic Forms in Normal Variables II Non-Central Case", **38**, (1967, AFOSR 67-0406, AD647-400).

THEORETICAL FOUNDATIONS FOR ASSOCIATIVE RETRIEVAL

Edward C. Bryant

Westat Research Analysis, Inc., Denver, Colorado

Project Description

One of the most difficult problems associated with information retrieval is that people tend to choose substantially different keys to describe a document. The development of an interface between the indexing and searching operations that allows latitude in the choice of indexing keys is necessary. The association of keys derived by the indexing operation to those of the search operation can be described as a linear transformation. This effort is to develop the underlying mathematical foundation for this association. Two basic questions will be addressed and explored for solution--1. What mathematical conditions or restrictions are imposed on the transformation? 2. How does one construct a transformation so that one has an optimized operation? A previous research effort has partially developed the mathematical concepts, but further work is needed to determine the constraints that allow these concepts to be valid.

Summary of Recent Research Results

Associative adjustments to a document file were considered as a means for improving retrieval. This study included the definition and theoretical investigation of the statistical properties of a generalized mismatch measure. Improvements in retrieval resulting from performing associative regression adjustments on a data file were examined both from the theoretical and experimental point of view. The expected gain in mismatch was presented in the final report as a function of various measurable characteristics of the file, such as error rates in indexing and the probability distributions of the associative adjustment criteria. Query adjustments using negative as well as positive correlations are considered and found to be ineffective. In a . limited size Patent Office file with a low indexing error rate, experimental results are presented applying (a) no associative correction, (b) the generalized mismatch with no associative correction, (c) associative correction, and (d) query adjustment. In general, the

results using the ordinary mismatch with an associative adjustment are superior to those using the more generalized quadratic mismatch or the query adjustment scheme.

Reports and Publications

Edward C. Bryant, D.T. Searls and R.H. Shumway, "Some Theoretical Aspects of the Improvement of Document Screening by Associative Transformations" (Final Technical Report, November 1965; AFOSR 66-0171, AD 628191).

Edward C. Bryant, D.T. Searls and D.G. Weinman, "Associative Adjustments to Reduce Errors in Document Screening" (Final Technical Report, March 1967; AFOSR 67-0980, AD 651630).

Section 3

TRANSMISSION OF INFORMATION

Many mechanisms exist by which information is transmitted. A simple organism can have a multiplicity of sensors for orientation in and adaption to its environment. Man has mechanisms within mechanisms compounded by his higher animal abilities of perception, awareness, and thought. Physical systems display order, a balancing of cause and effect, and adaption that suggest other mechanisms.

Living systems present a challenge to man's understanding of information processing. In packages smaller than he can fabricate, complex processes occur that he cannot yet model. What are the algorithms for seeing, for hearing, for remembering, for forgetting, for integrating one bit of information with another, for making quantum jumps in thought that lead to creative synthesis? How are impulses sensed, sorted, coded, transmitted, stored, recalled, evaluated? How are judgments made and decisions reached? How do organisms communicate with each other, and, in particular, how effectively does man communicate with man? Inquire into information transmission mechanisms ranges from studies of interactions at the subcellular level to those among men and between men and machines.

This task principally sponsors connecting research. However, other sponsors are not necessarily concerned with information science objectives that seek isolation of information-bearing parameters, models of information processing mechanisms, and techniques for improving communication processes. Since the range is broad and the budget is small, only a few projects can be supported. They highlight the difficulties that are inherent in attaining information science objectives.

59

3.1   A FORMAL ANALYSIS OF NATURAL AND ORAL VERBAL COMMUNICATION

Allen Barton and Emanuel A. Schegloff

Columbia University, New York, New York

Contract Number* AF 49(638)-1761

Project Description

No procedural description exists of naturally occurring social behavior or patterns of communication and information transmission that are used by members of society in normal or emergency situations. This effort will examine telephone call recordings between police and members of the public preceding and during a disaster situation and during a normal period toward developing a model of serial verbal communication interactions in natural conversation. The analysis is expected to provide formal descriptions of behavior in terms of the procedures (laws or rules) by which it is produced, and identify the classes or categories of "objects" (e.g., persons, situations, actions, etc.) that members of society, in effect, structure. Development of the descriptions will be interactive with the development of a language in which to construct the model. The research is expected to add to an understanding of: (a) the information value in the communications of lay monitors to professional or technical receivers; (b) relations between the legal system and other sectors of society; (c) how people perceive their involvement with events of social import; and (d) the ways in which the structure of a conversation mediates and shapes decisions and individual actions.

Summary of Recent Research Results

Attention has been concentrated on the very beginnings of the conversations taken to date. Several orders of phenomena and attendant analytic problems have been uncovered and described, and preliminary formulations of solutions have been developed. These can be grouped under three headings. 1) Sequencing phenomena: a set of formal properties describing the unit "summons-answer sequences" and its components have been elaborated to provide the elements of an adequate account of a standardized mode of initiating conversational interactions. 2) Categorization phenomena: a

preliminary formulation of a solution to a "first person categorization problem" has been worked out, relating the availability of a solution to the availability of a method for generating intended encounters. 3) Categorization transformation phenomena: two procedures for transforming the categorizations of parties to an interaction are described, and occasion is thereby found for explicating the operation of conversational devices earlier located in other contexts, such as the "correction-invitation device". These matters are discussed in the works listed in the appended bibliography.

Reports and Publications

Emanuel A. Schegloff,"The First Five Seconds: The Order of Conversational Openings" (Ph. D. Dissertation, Univ. of Calif. at Berkeley, Dept. of Sociology, 1967).

Emanuel A. Schegloff, "Sequencing in Conversational Openings", to be published in *Contributions in Ethnomethodology*, H. Garfinkel and H. Sacks, Eds. (Ind. Univ. Press, Bloomington, Indiana); to be reprinted in *Directions in Socioquistics: The Ethnography of Communication*, J. Gumperz and D. Hymes, Eds. (Holt, Rinehart and Winston).

3.2 **INTEGRATION OF THEORY AND
EXPERIMENT INTO A UNIFIED
CONCEPT OF VISUAL PERCEPTION**

Heinz Von Foerster and Humberts P. Maturana

University of Georgia, Athens, Georgia

Project Description

The principal aim of this study is the integration of recent anatomical and neurophysiological findings of structure and function of the afferent fibers in vertebrates on various levels of the optic tract with the theoretical results obtained in the study of computing networks. The goal of the study is the achievement of a unified concept of visual perception. This appears to be an opportune time for such an approach, since the apparent convergence of these fields will provide the mathematical basis for a unifying nomenclature and an understanding of the sophistication of the physiological problems involved.

This study is a joint effort of the Catedra Biologia Universidad de Chile, Santiago, Chile, represented by Dr. Humberto Maturana and the Department of Physics of the University of Georgia, represented by Dr. Heinz Von Foerster.

Summary of Recent Research Results

During the past year, difficulties in establishing synchronized schedules delayed the joint effort which will commence in November 1967. A major paper, however, was written by one of the principal investigators who summarized the present state of the theory of computing neural nets.

Reports and Publications

Heinz Von Foerster, *Currents in Modern Biology*, "Computation in Neural Nets", **1**, pp. 47-93 (1967).

3.3   AUDITORY SIGNAL DETECTION, CORRELATION AND TRANSMISSION

James P. Egan

Indiana University, Bloomington, Indiana

Project Description

The theory of signal detectability provides a framework for the investigation of the parameters that affect the detection of a weak signal by human observers. This framework is being used in a continuing attempt to understand the auditory processing system. The major concern is the experimental analysis of masking level differences (MLDs). An MLD is the increase in the detectability of a signal that results when an appropriate interaural condition is changed. These investigations of MLDs involve both continuous and burst masking noise. Systematic comparisons of the binaural with the monaural system may lead to an increased understanding of those binaural processes which markedly increase the detectability of a signal masked by noise. Measures of the detectability of a signal in an unstructured listening task, similar to the vigilance situation, will also be investigated.

Summary of Recent Research Results

Several investigations concerning masking level differences (MLDs) were recently accomplished. Under certain binaural conditions (MLD), signal detection is much better than under other conditions (non-MLD). It was found that when the MLD and the non-MLD conditions were presented in random order, detectability under each condition was unchanged. Furthermore, it was found that when the masking noise was presented as a burst, there was no effect on detection with a non-MLD condition. However, with an MLD condition, there was less improvement in detection with burst noise than with continuous noise. In addition, performance characteristics (ROCs) were investigated with MLD and non-MLD conditions. These functions were highly similar even though the signal energies required for equal detectability under these two conditions were markedly different. A study of both the detection and the recognition of weak signals showed that the listener did better than chance in recognizing one of two signals even when he said he did not detect it. Also, an extensive topical bibliography on signal detection theory and psychophysics was compiled.

63

Reports and Publications

James P. Egan, *J. Acoust. Soc. Am.*, "Masking-Level Differences as a Function of Interaural Disparities in Intensity of Signal and of Noise", **38,** No. 6, pp. 1043-1049 (Dec 1965; AFOSR 66-1341, AD640850).

James P. Egan and W. Benson, *J. Acoust. Soc. Am.*, "Lateralization of a Weak Signal Presented With Correlated and With Uncorrelated Noise", **40,** No. 1, pp.20-26 (July 1966; AFOSR 66-2675, AD643131).

James P. Egan, "Signal Detection Theory and Psychophysics: A Topical Bibliography", June 1967.

D.S. Emmerich, "Receiver Operating Characteristics Determined Under Several Interaural Conditions of Listening" (AFOSR 67-1561, AD655386).

Dennis McFadden, *J. Acoust. Soc. Am.*, "Masking-Level Differences with Continuous and With Burst Masking Noise", **40,** No. 6, pp. 1414-1419 (Dec. 1966; AFOSR 67-1498, AD654603).

Dennis McFadden, *J. Acoust. Soc. Am.*, "Detection of an In-Phase Signal with and without Uncertainty Regarding the Interaural Phase of the Masking Noise" (April 1967).

3.4   **CROSS-CULTURAL COMMUNICATION**
THROUGH NONVERBAL BEHAVIOR

Paul Ekman

Langley Porter Neuropsychiatric Institute,
San Francisco, California

Grant Number* AF-AFOSR-1229-67

Project Description

Nonverbal behavior can provide information about emotions,
traits, attitudes, personality, etc., but both sender and receiver usually
pay little explicit attention to this channel of communication. This
channel can be an important information source, both when the
nonverbal channel does not suffice (e.g., pilot gesticulating to crew)
as well as when the verbal channel may be used as a blind. Nonverbal
behavior, to date, has been less formalized and interpreted with less
precision than verbal behavior. This effort will examine measures for
nonverbal behavior and how the meaning of nonverbal behavior
varies among different cultures. The hypothesis will be tested that
certain types of nonverbal acts are universally understood, whereas
others have different communicative significance among different
peoples. Primitive peoples will be studied to determine whether
nonverbal expressions are innate or learned and thus specific in their
meanings within a culture. Analyses will be made of filmed behavior
and of observer judgments of selected examples of nonverbal be-
havior. Analyses will employ several measures of observers' judge-
ments, a recently developed Systematic Classification and
Analysis of Nonverbal Behavior (SCAN), and other methods to be
formulated during the research.

Summary of Recent Research Results

The first year of this project was devoted to planning and hypo-
thesis formation. A theoretical scheme has been devised which dis-
tinguishes five types of nonverbal behavior (emblems, illustrators,
regulators, affect displays, adaptors) in terms of their origin, usage
and coding. Specific predictions about universality or specificity
to particular cultures have been developed for each type of non-
verbal behavior. Pilot studies have been conducted which will be
further pursued including studies of primitive behavior in New
Guinea, of urban peoples in Japan and Argentina.

65

Reports and Publications

P. Ekman and W. Friesen, "Nonverbal Leakage and Deception Clues" (Speech to Langley Porter Neuropsychiatric Institute, San Francisco, April, 1967).

P. Ekman and W. Friesen, "Nonverbal Leakage" (Western Psychological Association Convention, 1967).

P. Ekman, W. Friesen and T. Taussig, "VID-R and SCAN: Tools and Methods in the Analyses of Facial Expression and Body Movement" (National Conference on Content Analysis, University of Pennsylvania, November, 1967).

P. Ekman and W. Friesen, "Origins, Coding and Usage: The Basis for 5 Categories of Nonverbal Behavior"(Conference on Communication Theory and Linguistic Models in the Social Sciences, Buenos, Aires, Argentina, October, 1967).

P. Ekman and W. Friesen, "Nonverbal Communication" in *Handbook of Communication*, Eds. Schramm, Parker, Maccoby, Pool, Frey and Fein (Rand McNally and Company, Chicago, Illinois; to be published in 1968).

3.5 MODEL AND MEASURES FOR THE HUMAN COMMUNICATION PROCESS

Forrest L. Brissey

University of Oregon, Eugene, Oregon

Project Description

Reduction of the processes of human communication to con-
trolled, quantitative study is essential to establish bases for effective
communication in systems where efficient transmission of infor-
mation is a criterion. This effort is exploring the measurement of
communicative effectiveness directly in terms of the non-verbal
decision behavior of both the source and the receiver. Research will
extend previous results on a conceptual model of human communi-
cation. A quantitative measure is sought of the degree to which dis-
plays are and may be efficiently and accurately encoded in a speci-
fied symbol system. The relevance of concepts of uncertainty and
redundancy as they have been derived from information theory and
applied to a linear sequence of binary events will be examined
toward the development of a useful measure of encodability for a
matrix of binary events. Empirical study will include measurement
of such factors as confrontation, status and decoding aptitude on
message-mediated decisions.

Summary of Recent Research Results

Current theoretical-conceptual work is concerned with relating
the concepts of *expectation, complexity, structure,* and *utility* to
the effectiveness of information transmission. A distinction is made
between primary and semiotic information. In part, the effort is
given to analyzing differences between systems in which semiotic
information is volunteered and those in which it is solicited.

Individual differences in the ability to estimate stimulus com-
plexity and in expressed preference for stimulus complexity are
theoretically important determinants of communicative effectiveness.
Several investigations have been undertaken in this area. Preliminary
analyses indicate stable over-estimation of complexity for low and
moderate levels of complexity and stable under-estimation for
high-complexity levels. A study to measure the effect of exposure
time on preference for complexity indicates that exposure times
between two and ten seconds yield highly similar, 'inverted U'
distributions for visually perceived displays.

67

The theoretical work mentioned above and experimental studies which are theory-relevant and yield significant results are being employed in the development of a system for communication training. The model is currently under both laboratory and field test.

Reports and Publications

Forrest L. Brissey, "An Experimental Technique for the Study of Human Communication" (Final Report, Sept 1964; AFOSR 64-2409, AD609946).

3.6   ASYNCHRONOUS AUTOMATA
AND DISCRETE-TIME SYSTEMS

George A. Bekey and Robert B. McGhee

University of Southern California,
Los Angeles, California

Project Description

Machine performance of complex information transfer and control operations depends on the development of suitable descriptive models. This effort is exploring models capable of describing asynchronous discrete-data systems. Necessary conditions were established for synthesizing optimum discrete systems based on pulse width and amplitude modulation. Proposed research will extend to other classes of combined modulation systems capable of describing, for example, information transmission in the neuromuscular system. An artificial quadruped was demonstrated whose controller for limb coordination was entirely a finite-state device. Further study is planned on problems pertaining to the prediction of stability, both practical stability relating to continuous state variables of the system and sequential stability associated with discrete aspects. Investigation of the problem of encoding large flow tables for asynchronous machines with code words of minimum length having the property that no critical races remain unsolved yielded a solution for tables of eight or less rows. Extensions of this and another approach to minimal coding will be examined.

Summary of Recent Research Results

Six specific research accomplishments are as follows:

1. An artificial quadruped controlled by a small special purpose digital computer has been successfully tested. The results obtained are in agreement with the performance predicted by a concurrently developed mathematical theory of legged locomotion.

2. Some new theoretical results on the transient behavior of combinational switching networks have been obtained. The results of this research both simplify the previously available theory and provide new algorithms for the detection of hazards in networks.

69

3. A new binary code for asynchronous sequential machines has been developed. This code permits the construction of machines with up to sixteen interval states using shorter code words than any previously available code.

4. Necessary conditions for optimal control of discrete-time systems with combined pulse-width and pulse-amplitude modulation were developed. These conditions can be used to select a sequence of control pulses from a wide class of modulator outputs.

5. An iterative technique for the hybrid computer solution of integral equations has been developed. This technique leads to substantial reductions in solution time, as compared to all-digital computation.

6. A new formulation of the signal flow graph technique has been developed for the study of discrete-data systems. It is shown that for continuous systems the new formulation reduces to the Mason formula.

Reports and Publications

George A. Bekey and E. A. Angel, "Asynchronous Finite State Models of Manual Control Systems" (U.S.C. Electrical Engineering Department Report No. 160, March, 1966).

George A. Bekey, R. Tomovic and J. C. Maloney, "Solution of Integral Equations by Hybrid Computation" (U.S.C. Electrical Engineering Department Report No. 178, April 1967; AFOSR; 67-1485 AD654633).

Robert B. McGhee, "Finite State Control of Quadruped Locomotion", in *Proceedings of the International Symposium on External Control of Human Extremities*, Dubrovnik, Yugoslavia, 1966 (Also appeared in *Simulation*, August 1967).

Robert B. McGhee, "Some Finite State Aspects of Legged Locomotion", (to be published in *Mathematical Biosciences*, 1967).

W. S. Meisel, "Hazards in Asynchronous Sequential Circuits" (to appear in *IEEE Transactions on Electronic Computers*).

W. S. Meisel, "A Numerical Analysis formula Useful in Fourier Analysis" (to appear in *Communications of the ACM*).

L. R. Nardizzi, "Analysis and Synthesis of Discrete-Time Systems With Control Signals of Variable Amplitude and Pulse-Width" (U.S.C. Electrical Engineering Department Report No. 205 June, 1967; AFOSR 67-1487, AD654632).

M. Sedlar and George A. Bekey, *IEEE Trans. Auto. Control*, "Signal Flow Graphs of Sampled-Data Systems - A New Formulation", AC-12, No. 2, pp. 154-161 (April 1967; AFOSR 67-1866, AD-656980).

70

3.7   ASSESSMENT OF PROGRESS IN CODING THEORY IN THE USSR

William H. Kautz

Stanford Research Institute, Menlo Park, California

Project Description

In developing methods for conveying information from one
source to another, formidable problems exist that pertain to a re-
quirement for transmitting and receiving with minimal error. Infor-
mation processing in computing machines must be error free. Machine
processing also raises questions of coding for data storage, reduction,
and retrieval, for the preparation of detailed repertoires of machine
commands, and for error detection and correction. Coding theory
developments may also furnish explanations for information trans-
mission in living systems and for errors that cause cells to go out of
control. Current activity in the U.S. and the USSR in coding theory
warrants a survey of Soviet work to enable continued informed re-
search in this field in the U.S. This effort will provide a state-of-the-
art assessment of the Soviet literature that will include papers not
previously adequately translated or distributed or described in
Western terminology. The review will be in two parts, for engineers
with limited detailed knowledge of coding theory, and for coding
specialists.

Summary of Recent Research Results

A survey of the available Soviet literature on coding theory and
related subjects is near completion. Approximately 300 papers have
been studied, of which 80 percent are adequately translated. Soviet
coding theorists have been active in essentially the same general areas
of pursuit as those in this country, namely, the synthesis, implemen-
tation, and evaluation of codes for various communication channels,
and for the transfer and processing of data in computers. While
considerable emphasis has been directed in this country to the syn-
thesis of general code families (e.g. Bose-Chaudhuri codes) and to
the description of general decoding techniques (e.g. threshold decod-
ing), many Soviet theorists have concentrated on seeking optimum
codes and decoding techniques for very specific codes. The Soviet
literature contains descriptions of previously unknown optimum
block codes of relatively short block length, and also descriptions of
novel decoding implementations for several of these codes. In addi-
tion the Soviet literature contains specific contributions on (1)
coding for the totally asymmetric channel. (2) coding for the

71

noiseless channel, (3) coding for the Gaussian channel, (4) coding for the feedback channel, (5) tight upper and lower bounds on code redundancy, and (6) state assignment for automata which correct transient errors.

72

Section 4

ADAPTIVE AND SELF-ORGANIZING SYSTEMS

Man, the information processor, intrigues man, the circuit designer, and man, the model builder. Man's nervous system is a fascinating structure, its operation a subject for conjecture. Some of man's sensory systems, his eye, his ear, are better understood and offer existence proofs to the engineer. Beyond what man can see are processes he calls perception, consciousness, and thought. Will he be able to find neural networks or molecules he can associate with these processes? Where and how is experience stored to provide for memory and recall?

Man and lower living systems embody principles of structure and performance which far surpass equations that he has been able to formulate and hardware he has been able to build. Man has repeatedly built machines exceeding his own physical capabilities. But, except to minor degrees, he has not yet built devices he can work with as extensions of himself in the way that he can work with other men. He has not built machines he can send into hostile and adverse environments that can report back to him selectively and reliably the information he would want to know. He has built machines which record vast quantities of data about the physical universe, yet few machines can distinguish the significant bits from the rest.

In information science terms, research to understand and model the dynamic processes of living systems is variously termed artificial intelligence, the field of adaptive and self-organizing systems, and when the construction of hardware is involved, bionics. The term cybernetics is sometimes used, because these systems incorporate control and feedback mechanisms that adjust their performance toward specific goals.

Research on adaptive systems may produce spectacular results over the short range, but such results can only be a first and gross approximation to what the long range holds. The mechanical eye or hand, the computer that voices the words "Bravo!" or "Come again?", the tracker that finds a target embedded in noise go part of the way. Mechanical devices with more and higher level features of human intelligence and adaption could teach and train man. Machines might not only help accelerate learning, but could also be used to stimulate and enhance man's inventiveness and creativity if more were known about his memory and thought processes. The equations by which man discriminates, makes decisions, integrates disjoint bits into a whole greater than the sum of its parts can be successively

73

approximated. It would be shortsightedness, however, to consider first approximations as more than precisely that.

Engineers, mathematicians, and perhaps administrators, tend to get carried away with the potential prospects of the intelligent machine. It is error to ignore data bases. The neurophysiologist knows something about the nervous system, but he has a lot more to discover. The psychiatrist and psychologist know something about behavior, but only a little something. The molecular biologist knows compound composition and structure, but what is the explanation for memory? Optimum results over the long range can only come from close cooperation across the range of disciplines that contribute pieces to the puzzle. The generalist and the specialist are both needed in the various fields, and they must be able to exchange ideas. Science is becoming one again as it simultaneously increases in specialization.

74

4.1 DESIGN OF A MACHINE TO EXPERIENCE SIMPLE SENSE DATA

F. Nowell Jones and James T. Culbertson

University of California, Los Angeles, California

Project Description

Sensory devices transform sensed data from one form into another. In this process some information is lost. In environments denied to humans (i.e., unmanned space probes) we must rely upon depleted information from sensory devices. If an automaton can be devised which will replicate human perceptual actions then this depletion of information can be removed. This effort is to further refine the theory that sense data in neural networks can be replicated in an automaton. This theory is based upon the use of a spacetime subspace in which the neural impulses over time form a network. This network is not only a description of the perception space but also includes the sense data and sense qualities. For non-trivial cases, the neuron must transfer energy with a minimum of scattering and the impulse must have a definable interconnection pattern. The first phase of the research will study these non-trivial cases to develop the minimization criteria and to define the requirements for interconnection. Following this effort an automaton (neural network model) will be designed. The design of this network will eventually allow for experimental verification of this theory.

Summary of Recent Research Results

The derivation of a perception space from the network of neuron impulses over time has been completed. This derivation includes representing such qualities as color and pain sensations. The corresponding hardware would be a paralyzed (no output) automaton that would experience these sense data. Since March 1967, the main problem has been to simplify the theoretical design so that sense data experienced by the automaton can also be experienced by an observer. The corresponding hardware would require neural connections from the automaton to the nervous system of the observer. The automaton would then be an annex to the brain of the observer, provided the scattering is minimized at the automaton-observer interface. The present design would allow for these first attempts at verification.

4.2   TRANSLATION AND FIRST-STAGE FORMAL ANALYSIS OF NATURAL CONVERSATION IN THAILAND

Michael Moerman

University of California, Los Angeles, California

Current Grant Number* AF-AFOSR-1167-66

Project Description

Several significant objectives are inherent in the analysis of natural languages of the Orient: contribution to the methodology being developed for analyzing natural language to determine how to represent its structure and how it is used for communication and to convey meaning; obtaining an understanding of the behavior and modes of communication of a people for effective cross-cultural interaction; and the enrichment of ethnographic knowledge. This study is for the transcription, translation, and preliminary analysis of a unique corpus of Tai-Lue natural conversation. The analysis will explore the influence of situation, term, reputation, and social control on communication patterns. The analysis will also examine the effect of subject (e.g., officials, road building, fertilizer, missionaries, Vietnam) and directed as distinguished from natural conversation on the language used. The research may also suggest scientific approaches to the observation and analysis of natural social events.

Summary of Recent Research Results

Two preliminary indications encourage belief in the possibility of doing ethnography through direct observation of natural conversation.

1. Culturally typical ascriptions of action and of attitude to culturally recognized persons can be discovered by means of determining the referents of pronouns, pro-verbs, and deleted nouns and verbs.

2. It seems possible to investigate interpersonal influence and dominance through a corpus of natural conversations and without recourse to such distorting external artifacts as questionnaires, genealogies, etc....

The utterance tieing rules through which Lue conversations are made coherent to their participants (e.g., question/answer rules, command/response, PRO-forms) promise to be significantly similar to those governing English conversation (4,5).

Reports and Publications

Michael Moerman, *American Anthropologist,* "Reply to Narroll" (to be published).

Michael Moerman, "A Little Knowledge", in *Contributions in Ethnomethodology,* H Garfinkel and H. Sacks, eds. (University of Indiana Press, to be published).

Michael Moerman, "Being Lue: Uses and Abuses of Ethnicity", in *Proceedings of the 1967 American Ethnological Society Meeting* (to be published).

Harvey Sacks, Unpublished Lectures, 1965-1967.

Emanuel Scheglof, "The First Five Seconds" (Ph. D Dissertation, University of California at Berkeley, 1967).

RESEARCH IN INFORMATION PROCESSING AND COMPUTER SCIENCE

Allen Newell and Alan J. Perlis

Carnegie Mellon University (formerly
Carnegie Institute of Technology),
Pittsburgh, Pennsylvania

Contract Numbers* SD-146 F44620-67-C-0058

Project Description

This effort supports fundamental research in Computer Science and Information Processing with the goal of achieving deep understanding and concise formalization of the basic concepts, techniques, and processes in the following areas: 1) man-machine communication, 2) programming systems, 3) theory programming, 4) problem-solving, 5) hardware and software system design, 6) language design and development, 7) communication systems, 8) computer networks, 9) automata theory and logic, 10) numerical analysis, and 11) linguistics. Specific research goals within these broad areas include the development of specialized programming languages for such purposes as system design, simulation, compiler building, and symbol manipulation. An attempt is being made to develop a verifying compiler which would insure that logical errors do not exist in a program compiled and accepted by this compiler. Research on general aspects of time-sharing methodology, along with an analysis of the impact of time-sharing on language design, is being undertaken. A study of the engineering problems associated with the development of computer networks will also be made. Problem-solving research is aimed at investigating the reasoning mechanisms necessary to solve complex problems. Methods of problem representation are sought that will be independent of the mechanism chosen to code the solution algorithm for a given problem. A study of the application of graphic development to the field of information processing will be undertaken toward building a strong graphic capability. This capability will aid to advance information processing concepts and techniques in areas such as design, computer-aided instruction, and information handling in general.

Summary of Recent Research Results

FORMULA ALGOL has long been under development at Carnegie-Mellon University (CMU) as an extension to Algol in that it incorporates symbol manipulation in an integrated way. Besides the FORMULA ALGOL system itself, which has finally become running and documented (though still experimental) several major studies have come to fruition. These studies both use Formula Algol and derive their inspirations from it. **Tim Standish's** research involves a further extension called a Data Definition Facility in which one specifies (in a way similar to Backus Normal Form) the data structures one wants to manipulate. The system then provides a notation and the facility for manipulating these objects. **Renato Iturriaga** has developed a set of algorithms to determine the convergence of infinite series and the values of limits. Limiting processes and infinite series are everyday working tools of the mathematical analyst and a thorough assimilation of them into symbol manipulation is an essential part of developing an automatic mathematical assistant. **Robert Caviness** studied the representations of formula expressions in a way that would give meaning to the so-called simplification problem. As a step toward this goal, he has defined the concepts of canonical and normal forms as alternatives to the controversial and ill-defined concept of a simplified form. Finally, **Steven Coles** has used FORMULA ALGOL as the basic language in constructing a graphic system which accepts simple geometric figures drawn on the SCOPES (Carnegie-Mellon's graphic hardware display system) plus sentences about the diagram in quasi-English. His system GRANIS then determines whether the sentences are true. The system has been limited to a simple demonstration system due to the restraints on the G-21 memory.

Carnegie-Mellon has acquired an IBM 360/67 and has thus been centrally involved in bringing into being a large third generation time-sharing system. Very early in the development of the 67, we determined that the swapping drum was one of the key limiting features in the performance of the system as currently designed. As an alternative we have made use of a million words, of *large core storage* and have recently written a paper on the basic issue of memory swapping (24). The Center is now actively engaged in bringing up TSS on the 360/67.

Additional programming activities have involved the development of a version of L[6] (Low Level Linked List Language, by K. Knowleton of Bell Labs) for the 360; our system is called *1. We have finished the detailed specifications of a compiler-compiler system called CABAL which is fashioned upon the experience gained with an earlier compiler system FSL (Formal Semantic Language). FSL was used extensively in the coding of FORMULA ALGOL.

Robert Floyd's research includes an adequate and formal basis for proving matters of equivalence, correctness and termination of programs. A promising theoretical base has been obtained (an

79

article is currently in press (17) and work is proceeding on a programming system named the "Verifying Compiler".

Work in problem solving in the last year, (i.e., subsequent to the conclusion of a major effort on GPS presented at the 1967 Spring Joint Computer Conference) has been mostly concerned with the study of human problem solving. Here the problem has been to develop techniques for extracting the behavioral regularities from the raw record. A detailed study (29) has been completed by Allen Newell and a report is in press (28).

Reports

1. Gordon C. Bell, "Timed-Shared Computers" (Carnegie Institute of Technology Report, May, 1967; AFOSR 67-1618, AD-655380).
2. R. T. Braden and Alan J. Perlis, "An Introduction in Computer Programming" (Carnegie Institute of Technology Report, June 1965).
3. Stephen L. Coles, "Syntax Directed Interpretation of Natural Language" (CIT Report, May, 1967; AFOSR 67-1758, AD-655923).
4. David C. Cooper, "The Equivalence of Certain Computations" (CIT Report, 1966; AFOSR 67-0253, AD467634).
5. David C. Cooper, "Computer Programs and Graph Transformations" (CIT Report, September 1966).
6. R. Dupchak, "TIPL: Teach Information Processing Language" (Rand Corporation Report, Santa Monica, California, October 1963).
7. R. Dupchak, "LIPL: Linear Information Processing Language" (RAND Corporation Report, Santa Monica, California, February 1965).
8. Jay Early, "Generating a Recognizer for a BNF Grammar" (CIT Report, June 1965).
9. Jay Earley, "FORMULA ALGOL Manual" (CIT Report, June 1967).
10. George W. Ernst and Allen Newell, "Generality and GPS" (CIT Report, January 1967; AFOSR 67-0756, AD809354).
11. Arthur Evans Jr., "Syntax Analysis by a Production Language" (CIT Report; AFOSR 67-0254, AD625465).
12. Jerome A. Feldman, "A Formal Semantics for Computer Oriented Languages" (CIT Report, May 1964).
13. J. W. Fierst (Ed.), D.M. Blocker, R. T. Braden, A. J. Evans and R. B. Grove, "ALGOL-20 · A Language Manual" (CIT Report, February 1965).
14. N. V. Findler, "An Information Processing Theory of Human Decision-Making Under Uncertainty and Risk (CIT Report, July 1964; AFOSR 67-2129, AD658486).
15. Robert W. Floyd, "Non-Deterministic Algorithms" (CIT Report, November 1966).

16. Robert W. Floyd, "New Proofs of Old Theorems in Logic and Formal Linguistics" (CIT Report, November 1966; AFOSR 67-2517, AD660886).
17. Robert W. Floyd, "Assigning Meanings to Programs", to be published.
18 Abraham Ginsburg, 'Six Lectures on Algebraic Theory of Automata" (CIT Report, November 1966).
19. Michael M. Gold and Lee L. Selwyn, "Toward Economical Remote Computer Access" (CUT Report, July 1967; AFOSR 67-2018, AD657783);
20. G.J. Hansen "SOL-20" (CIT Report, April 1965; AFOSR 67-2516, AD660885).
21. R. Iturriaga, T. A. Standish, R. Krutar and J. Earley, "The Implementation of Formula ALGOL in FSL" (CIT Report, October 1966).
22. Sheldon Klein, S. L. Lieman and G. E. Lindstrom, "Diseminer: A Distributional-Semantics Inference Maker" (CIT Report, June, 1966; AGOSR 67-0255, AD636380).
23. R. Kurki-Suonio, "On Character Set Reduction" (CIT Report, August 1965).
24. Hugh C. Lauer, "Bulk Core in a 360/67 Time-Sharing System", System", in *Proceedings of the 1967 Fall Joint Computer Conference,* Anaheim, California (Thompson Books, Washington, D.C., 1967; AFOSR 67-1968, AD657782).
25. Ralph L. London, 'A Computer Progam for Discovering and Proving Sequential Recognition Rules for Well-Formed Formulas Defined by A. Backus Normal Form Grammar" (CIT Report, May 1964; AFOSR 67-0259, AD804036).
26. Allen Newell and Herbert A. Simon, "An Example of Human Chess Play in the Light of Chess Playing Programs" (CIT Report, August 1964; AFOSR 67-0206, AD619386).
27. Allen Newell, Jay Earley and Fred Haney, "*1 Manual" (CIT Report, June 1967).
28. Allen Newell, "One the Analysis of Human Problem-Solving Protocols" (CIT Report, July 1967).
29. Allen Newell, "Studies in Problem Solving: Subject 3 on The Crypt-Arithmetic Task DONALD + GERALD = ROBERT (CIT Report, July 1967; AFOSR 67-2169, AD658485).
30. Joyce Nissenson (ed.), "Computer Science Research Review" (CIT Report, published in 1967; AFOSR 67-0252, AD645294).
31. Daivd L. Parnas, "System Function Description ALGOL" (CIT Report, February 1965).
32. David L. Parnas, "Sequential Equivalents of Parallel Processes" (CIT Report, February 1967; AFOSR 67-0755, AD809415).
33. Alan J. Perlis, R. Iturriaga and T. A. Standish, A Preliminary Sketch of Formula ALGOL (CIT Report, July 1965; AFOSR 67-2209, AD659156).
34. Alan J. Perlis, R. Iturriaga and T. A. Standish, "A Definition of Formula ALGOL (CIT Report, March 1966).
35. Jessie T. Quatse, "Design of the G-21 Multi-Processor System (CIT Report, February 1965; AFOSR 67-0257, AD804037

36. Jessie T. Quatse, "A Visual Display System Suitable for Time-Shared Use" (CIT Report, June 1965).
37. D. Ross, "MTHAT Assembler for the CDC G-21" (CIT Report, July 1966; AFOSR 67-0256, AD803897).
38. T. A. Standish, "A Data Definition Facility for Programming Languages" (CIT Report, May 1967; AFOSR 67-2045, AD658042).
39. E. Stefferud, "The Logic Theory Machine: A Model Heuristic Program" (RAND Corporation, Santa Monica, California, 1963).
40. J.C. Strauss and W.L. Gilbert, "SCADS: A Programming System for thee Simulation of Combined Analog Digital Systems" (CIT Report, March 1964; AFOSR 67-2518, AD660887).
41. T. G. Williams, "Some Studies in Game Playing with a Digital Computer" (CIT Report, July 1965; AFOSR 67-0258, AD634821).

Publications

Gordon C. Bell and H. W. Pirtle, "Time Sharing Bibliography", in *Proceedings of the IEEE*, **54**, number 12 (December 1966) pp. 1764-1766.

Gordon C. Bell, "Communication and Computers", in *Carnegie Review*, number 12 (July 1967) pp. 6-15.

P.G. Ciarlet, M.H. Schultz and R.S. Varga, "Numerical Methods of High-Order Accuracy for Nonlinear Two-Point Boundary Value Problems", in *Proceedings of the International Colloquim C. N. R. S.* (Besancon, France, 1966).

P.G. Ciarlet, H.G. Schultz and R.S. Varga, *Numerische Mathematik*, "Numerical Methods of High-Order Accuracy for Nonlinear Boundary Value Problems 1. One Dimensional Problems", **9**, pp. 394-430 (1967).

David C. Cooper, "Theorem Proving in Computers", in *Advances in Programming and Non-Numerical Computation*, L. Fox. ed. (Permagon Press, 1966) pp. 155-182.

David C. Cooper, *The Computer Journal*, "The Equivalence of Certain Computations", **9**, number 4, pp. 45-52 (1966).

David C. Cooper, "Mathematical Proofs about Computer Programs", in *The Machine Intelligence*, D. Miche, ed. (Oliver and Boyd, November 1966).

David C. Cooper, "Reduction of Programs to a Standard Form by Graph Transformations", in *Proceedings of International Seminar on Graph Theory and Its Applications*, Rome Italy, July 1966 (January 1967).

David C. Cooper, "Some Transformations and Standard Forms of Graphs with Applications to Computer Programs", in *Machine Intelligence II*, D. Miche, ed. (Oliver and Boyd, June 1967).

M.C. Cramer and J. C. Strauss, "A Hybrid Oriented Interactive Language", in *Proceedings of the Twenty-first National*

82

Conference of the ACM (Thompson Book Company, Washington D.C. August 1966).pp. 479-498.

George Ernst and Allen Newell, "Some Issues of Representation in a General Problem Solver", in *Proceedings of the Spring Joint Computer Conference*, Atlantic City, N J., 1967, pp. 583-600.

S. Even and A.R. Meyer, *IEEE Transactions on Electronic Computers*, "Test for Planarity of a Circuit Given by an Expression", **EC-15**, number 3, pp. 372-375 (June 1966).

P.C. Fischer, A. R.Meyer and A. L. Rosenberg, "Real-Time Counter Machines", in *Conference Record of 8th Annual Symposium on Switching and Automata Theory*, IEEE Computer Group (October 1967) pp. 148-154.

Robert W. Floyd, ed., *Communications of the ACM*, "Proceedings of the Symposium of Symbolic and Algebraic Manipulation" (August 1966).

Frank Hole and Mary Shaw, "Computer Analysis of Chronological Seriation", *Rice Unversity Studies*, **53**, number 3 (Summer 1967).

R. Iturriaga, T. A. Standish, R. Krutar and J. Earley, "Techniques and Advantages of Using the Formal Compiler Writing System FSL to Write a Formula ALGOL Compiler", in *Proceedings of the AFIPS Spring Joint Computer Conference*, Boston, April 26-28, 1966 (Spartan Books, Washington D.C., 1966) pp. 241-252.

M. Lees and M.H. Schutz, "A Leray-Schauder Principle for A-Compact Mappings and the Numerical Solution of Non-Linear Two-Point Boundary Value Problems", in *Numerical Solutions of Nonlinear Differential Equations*, Donald Greenpan, ed. (J. Wiley and Sons, Inc,. 1966).

L. E. Liphin, W. Watt and R. S. Kersch, *Annals of the New York Academy of Sciences*, "The Analysis Synthesis and Description of Biological Images, **128**, pp. 984-1012 (January 1966).

D. W. Loveland, *Zeitschrift fur Mathematik Logik and Grundlagen Mathematik*, "A New Interpretation of the Von Mises' Concept of Random Sequences", **12**, number 4 (1966).

D.W. Loveland, *Trans. of the American Mathematical Society*, "The Kleene Hierarchy Classification of Recusively Random Sequences, **125**, number 3, pp. 497-510 (1966).

A.R. Meyer and D. M. Ritchie, "The Complexity of Loop Programs", in *Proceedings of the 22nd National Conference of the ACM* (Thompson Book Company, Washington D.C., August 1967) pp. 465-470.

A. R. Meyer, A. L. Rosenberg and P. C. Fischer, "Turing Machines with Several Read-Write Heads", in *Conference Record of 8th Annual Symposium on Switching and Automata Theory*, IEEE Computer Group (October 1967, pp. 117-127.

Allen Newell and Herbert A. Simon, "Information Processing in Computer and Man", in *Science in Progress*, W.R. Brode, ed. (Yale University Press, New Haven Conn., 1966) pp. 333-362.

Allen Newell, "Discussion of Papers by Dr. Gagne and Dr. Hayes", in *Problem Solving: Research, Method, and Theory*, B. Kleinmuntz, ed. (J. Wiley and Sons, Inc., New York, 1966) pp. 171-182.

83

David H. Nickerson, "Managing a Computation Center", in *Computer Science Research Review* (Carnegie Institute of Technology, 1966) pp. 6-17.

David L. Parnas, *Communications of the ACM*, "A Language for Describing the Function of Synchronous Systems, **9**, number 2, pp. 72-76 (1966).

David L. Parnas, *Communications of the ACM*, "On Facilitating Parallel and Multi-Processing in ALGOL", **9**, number 4, p. 257 (1966).

David L. Parnas, *Communications of the ACM*, "On the Preliminary Report of C³S", **9**, number 4, pp. 242-243 (1966).

David L. Parnas, *Journal of Engineering Education*, "On the Use of the Computer in Engineering Education Without a Programming Prerequisite", **56**, number 8 (1966).

David L. Parnas, *Journal of the ACM*, "State Table Analysis of Programs in an ALGOL - like Language", **14**, number 1, pp. 1-9 (June 1967).

Alan J. Perlis, *Communications of the ACM*, "A New Policy for Algoithms", **9**, number 4, p. 255 (1966).

Alan J. Perlis, *Journal of the ACM*, "Synthesis of Algorithmic Systems", **14**, number 1, pp. 1-9 (June 1967).

Alan J. Perlis, *Communications of the ACM*, "A Proposal for Definitions in ALGOL" **10**, number 4, pp. 204-219 (April 1967).

Jessie T. Quatse, "Generality in Computer Design", in *Computer Science Research Review* (Carnegie Institute of Technology, pp. 53-60 (1966).

Jessie T. Quatse, "Time Shared Troubleshooter Repairs, Computers On-Line", in *Electronics* (McGraw Hill, January 1966) pp. 97-101.

Jessie T. Quatse, "Feedback Choke Reduces Power Supply Ripple", in *Electronics* (McGraw Hill, June 1966) p. 74.

Jessie T. Quatse, and Roy A. Kein, *IEEE Transactions of Electronic Computers*, "A Parallel Accumulator for a General Purpose Computer", **EC-16**, number 2, pp. 165-171 (April 1967).

M. H. Schultz, *Proceedings of the American Mathematical Society*, "A Generalization of the Lax-equivalence Theorem", **17**, pp. 1034-1035 (1966).

M. H. Schultz, *American Mathematical Society Bulletin*, "On the Stability of Discretizations", **72**, number 6, pp. 950-953 (November 1966).

M. H. Schultz, *Journal of Mathematics and Mech.*, "Convergence of Difference Approximations to Cauchy Problems in the Space of Tempered Distributions", **16**, number 10, pp. 1117-1130 (April 1967).

84

M. H. Schultz and R. S. Varga, *Numerische Mathematik*, "L-Splines", **10**, pp. 345-369 (1967).

Herbert A. Simon, "Reflections on Time Sharing from a User's Point of View", in *Computer Science Research Review*, Carnegie Institute of Technology, pp. 43-51 (1966).

Herbert A. Simon and Jeffrey M. Paige, "Cognitive Processes in Solving Algebra Word Problems", in *Problem Solving*, B. Kleinmuntz, ed. (J. Wiley and Sons, New York, 1966) pp. 51-119.

Herbert A. Simon, "Political Research: The Decision-Making Framework", in *Varieties of Political Theory*, D. Easton, ed. (Prentice-Hall, Englewood Cliffs, New Jersey, 1966) pp. 15-24.

Herbert A. Simon and George W. Baylor, "A Chess Mating Combinations Program", in *Proceedings of the Spring Joint Computer Conference*, Boston, April 26-28, 1966 (Spartan Books, Washington D.C., 1966) pp. 431-447.

Herbert A. Simon, "Thinking by Computers", in *Mind and Cosmos: Essays in Contemporary Science and Philosophy*, R. G. Colodny, ed. (University of Pittsburgh Press, 1966) pp. 3-21.

Herbert A. Simon, "Scientific Discovery and the Psychology of Problem Solving", in *Mind and Cosmos: Essays in Contemporary Science and Philosophy*, R. G. Colodny, ed. (University of Pittsburgh Press, 1966) pp. 22-40.

Herbert A. Simon, *The American Scholar*, "A Computer for Everyman", **35**, pp. 258-264 (Spring, 1966).

Herbert A. Simon, "The Impact of New Information Processing Technology: 1. On Managers; 2. On the Economy", in *Commercial Letter*, Canadian Imperial Bank of Commerce, Toronto (October 1966) pp. 1-12.

Herbert A. Simon, *Journal of Symbolic Logic*, "A Note on Almost-Everywhere Definability (Abstract), **31**, pp. 705-706 (December 1966).

Herbert A. Simon, *Psychometrika*, "A Note on Jost's Law and Exponential Forgetting", **31**, pp. 505-506 (December 1966).

Herbert A. Simon and Nicholas Rescher, *Philosophy of Science*, "Cause and Counterfactual", **33**, pp. 323-340 (December 1966).

Herbert A. Simon, *Psychological Review*, "Motivational and Emotional Controls of Cognition", **74**, pp. 29-39 (January 1967).

Herbert A. Simon, "Programs as Factors of Production", in *Proceedings of the 19th Annual Winter Meeting, 1966, Industrial Relations Research Association*, Madison, Wisconsin, pp. 178-188 (1967).

Herbert A. Simon, " The Changing Theory and Changing Practice of Public Administration", in *Contemporary Political Science*, Ithiel de Sola Pool, ed. (McGraw Hill, New York 1967) pp. 86-120.

Herbert A. Simon, *American Documentation*, Letter to the Editor: "Literary Output of Scientists and Engineers", **28**, pp. 113-114 (April 1967).

Herbert A. Simon, *Journal of Management Studies,* "The Business
  School: A Problem in Organizational Design", **4,** pp. 1-16
  (February 1967).

Herbert A. Simon, *Think,* "Information Can Be Managed", **33,**
  number 3, pp. 8-12 (May-June 1967).

Herbert A. Simon and Lee W. Gregg, *Journal of Mathematical
  Psychology,* "Process Models and Stochastic Theories of
  Simple Concept Formation", **4,** pp. 246-276 (June 1967).

Herbert A. Simon, *Educational Record,* "The Job of a College
  President", **48,** pp. 68-78 (Winter 1967).

J.C. Strauss, "Optimal Tracking of Nonlinear Dynamic Systems",
  in *Symposium on Optimization Techniques,* A. Lavi and T.
  Vogel, eds. (J. Wiley and Sons, Inc., New York, 1966) pp.
  449-469.

J. C. Strauss, "Basic Hytran Simulation Language" in *Proceedings
  of the AFIPS Fall Joint Computer Conference,* San Francisco,
  November 1966 (Spartan Books, Washington D.C., 1966) pp.
  603-611.

D. S. Wile, A. van Dam and R. G., Munck, "The Brown University
  Student Operating System", in *Proceedings of the Twenty-
  second National Conference of the ACM (Thompson* Book
  Company, Washington D.C., 1967) pp. 427-439.

4.4 TIME SHARED COMPUTING SYSTEMS

Thomas E. Kurtz, John G. Kemeny and Myron Tribus

Dartmouth College, Hanover, New Hampshire

Project Description

Time Sharing is the simultaneous shared use of a computer system by independent users expecting short or appropriate responses, within the limits of the request and system, to computational demand stimuli. This concept, its application and implementation, is a relatively recent one in computer science, and presents many challenges, unsolved problems, and future promise to a significant cross section of scientific disciplines. This research effort will investigate many aspects of the total time-sharing philosophy with the goal of extending basic concepts and theories associated with computation, in general and time-sharing, specifically. General areas of interest will include the following. 1) The design of advanced peripheral devices that can be produced economically and run in conjunction with a time-shared system. An important consideration in this area is to insure that such devices do not put an excessive load on the overall system. 2) The development of new languages and software to facilitate the man-machine interface and for dealing with special problems such as large file manipulation and computer-to computer communications. 3) The development of the computer system for use in various practical situations such as library management, teaching, and as a component in a computer network. A central area of interest will be the design and development of terminal systems and satellite computers. The development of simple, inexpensive graphical displays and sophisticated consoles, as well as special purpose satellite computers, that can function in a time shared environment, will contribute to the partnership of man and computer.

Summary of Recent Research Results

This AFOSR research effort received its initial sponsorship less than one year ago, and will be reported in the next issue of this publication.

4.5 COMPUTER SIMULATION OF NATURAL PATTERN GENERATION PROCESSES

Dan Cohen

Hebrew University, Jerusalem, Israel

Project Description

The growth and structural patterns of biological systems suggest the existence of levels of organization and information communication networks that could be exploited in electronic communication systems, if their coding patterns were understood. In this effort, the generation of particular biological patterns will be simulated by the application of ordered sets of generation rules. The rules will be modified and extended to increase the resemblance between the output of the simulation and the natural patterns, and alternative sets of rules will be explored to test their uniqueness in generating particular patterns. Simple patterns in two dimensions will be examined first. Results are expected to disclose such factors as

   a. The upper limit that can be set to the minimal amount of information required to specify a particular pattern,
   b. The most economic language by which a pattern can be classified and described,
   c. The minimal degree of complexity for a pattern,
   d. General requirements for the generation of classes of patterns.

Summary of Recent Research Results

Recently, research on this project has been directed toward finding models for the generation of patterns of cells from an initially random arrangement. Special emphasis is given to the generation of patterns in a growing mass of cells.

Reports and Publications

Dan Cohen, "Computer Simulation of Biological Pattern Generation: A Preliminary Report" in *Quarterly Progress Report Number 82* (Research Electronics Laboratory, Massachusetts Institute of Technology, July, 1966).

Dan Cohen, *Nature*, "Computer Simulation of Biological Pattern Generation Processes" (October 1967).

4.6   CRITERIA FOR ADAPTIVE SYSTEMS

Gerhard L. Hollander

Hollander Associates, Fullerton, California

Project Description

Research on adaptive systems has ranged from the formulation of
mathematical models to hardware design in attempts to describe and
construct systems and devices capable of automatic adjustment to
changing environmental conditions. A given research effort most often
concentrates on a particular hypothesized approach to adaptation,
often in terms of a unique notation that makes analysis of the ap-
proach difficult for quantitative evaluation and comparison. In this
effort, various approaches to adaptive system design and construction
are being examined toward the development of criteria for quantita-
tive evaluation. External characteristics of the various systems are
being studied from a functional viewpoint. Factors that are being
examined include range of adapting ability, convergence spe~d,
complexity of design, percentage of components required fo₁
learning ability, and flexibility. Results should disclose the funda-
mental characteristics of the various systems and enable their classi-
fication and comparison. The system-independent criteria should
permit an assessment of the various approaches for use in decisions
on their exploitation.

Summary of Recent Research Results

A review of over 100 research reports shows that efforts in
adaptive systems can be divided into three broad categories: the
network, heuristic, and servomechanism approaches. Common cri-
teria apply to all three, but numerical values differ significantly. The
servomechanism implementation has been analyzed further to test
the approach.

4.7 EVOLUTIONARY AND MULTILEVEL INFORMATION PROCESSING NETWORKS

Earl E. Gose

University of Illinois, Chicago, Illinois

Project Description

The availability of high-powered analog and digital machines
has opened a new area of research, the study of complex networks
exhibiting characteristics of adaptation and learning. This study re-
quires, first, the postulation of mathematical descriptions of the
performance of such networks, followed by computation of the
parametric possibilities. This effort is studying procedures for the
training of multilevel networks consisting of more than one layer of
adjustable parameters. Evolutionary processes are also being explored
toward the specification of optimum structures for multilevel net-
works. These procedures involve either the mutation or replacement
of network elements. Criteria must be selected for evaluating the
elements, and decision procedures are required for effecting changes
in the network structure. Preliminary results show that the evolution-
ary approach requires significantly fewer elements with an appreci-
able decrease of error over training procedures. An attempt will be
made to design an adaptive evolutionary network descriptive of
physical world situations in terms of nonlinear integro-differential
equations. Results should add to an understanding of requirements
for the specification of intelligent automata.

Summary of Recent Research Results

A theory has been derived for obtaining the optimal replacement
schedule for an evolving population of network elements. The theory
allows for variable quality distribution functions and training costs
as a function of the generation. Several element selection criteria
have been evaluated and compared. All of them led to improvement
over the non-evolutionary case.

A multilevel adaptive network of discrete logical elements has
been simulated and tested with the task of learning logical functions.
Learning has been demonstrated and a larger network is being pro-
grammed. Learning has also been demonstrated for a multilevel net-
work of continuous functions and the properties of its performance
space are being evaluated.

90

Reports and Publications

Earl E. Gose, *IEEE Trans. Electron. Computers,* "A Synthesis Technique for Networks Consisting of Logical Functions Feeding a Linear Summation Element", **vec 14,** no. 2, pp. 254-256 (April 1965; AFOSR 65-1590, AD623500).

Earl E. Gose, *Inform. Control,* "An Adaptive Network for Producing Real Functions of Binary Inputs", **8,** No. 2, pp. 111-123 (April 1965; AFOSR 65-2686, AD629836).

Earl E. Gose, "Improving the Population of Elements in a Pattern Recognition Network by Periodic Selection and Replacement", presented at the International Symposium on Information Theory, San Reno, Italy, 11-15 Sept. 1967.

A.H. Klopf, "Evolutionary Pattern Recognition Systems" (Supplementary Technical Report; AFOSR 66-1342, AD637492).

A.N. Mucciardi, "Adaptive Pattern Recognition Using Non-Linear Elements" (Supplementary Technical Report No. SRC 65-A-64-22, Oct. 1964; AFOSR 65-1701 AD623215).

A.N. Mucciardi and Earl E. Gose, *IEEE Trans. Electron. Computers,* Evolutionary Pattern Recognition in Incomplete Nonlinear Multithreshold Networks, **vec-15,** No. 2, pp. 257-261 (April, 1966).

4.8 THEORY AND APPLICATION OF COMPUTATIONAL PRINCIPLES IN COGNITIVE SYSTEMS

Heinz Von Foerster

University of Illinois, Urbana, Illinois

Project Description

"Cognitive processes" enable a system to acquire knowledge about the regularities, redundancies, and structures of its environment. Without such knowledge, symbolic discourse cannot occur and all exchanges of information must be over channels that transmit the full amount of information. This research is a multi-phase, multi-discipline effort toward specifying the hierarchies of transformations of information that lead from the perception of particulars to the manipulation of generalized symbolic representations of the perceived, and the inverse transformations that lead from general commands to specific actions and from general concepts to specific utterances that constitute cognitive processes. Proposed research will continue theoretical and applied studies and experimentation on modalities of inference, particularly with respect to non-linguistic representations of linguistic structures, theories of self-referential systems, extensions of cybernetic theory, particularly with respect to structural self-organization, evolution, and the behavior of large, complex systems, realizations of symbolic discourse between man and machines, properties of computing networks, the discovery and development of structural constraints for automata, and the organization and performance of pattern recognizers and processors.

Summary of Recent Research Results

On the experimental side of this research project, various electronic systems have been completed and are operational. Among these is the Visual Image Processor, a high speed parallel computer which processes 10^7 bits per frame by operating directly on neighborhoods of $1.6 \cdot 10^5$ reference points in two-dimensional images; a Speech Analyzer which uses redundancies in natural speech and maps "speech events" into unique significant subsequences of corresponding "machine events"; a Speech Synthesizer which utilizes the information transmitted in the second formant only; and an Automatic Stochastic Predictor which permits a systematic study of Pavlovian reflexology and related conditioning and learning situations.

Books and Articles

W.A. Ainsworth, *Science*, "Electrolytic Growth of Silver Dendrites, **146**, pp. 1294-1295 (1964).

A.M. Andrew, *Electronic Engr.*, "A Beam-Blanking Circuit", **38**, number 456. pp. 90-91 (1966).

A.M. Andrew, *Electronic Engr.*, "Counting to 1,099, 508, 482, 050 Without Carries", **38**, number 456, pp. 172-175 (1966).

A.M. Andrew, *Kybernetik*, "To Model or not to Model", **3**, pp. 272-275 (1967).

W.R. Ashby, H. von Foerster and C.C. Walker, "The Essential Instability of Systems with Threshold and Some Possible Applications to Psychiatry", in *Nerve, Brain, and Memory Models*, N. Wiener and J.P. Schade, eds. (Elsevier Press, Amsterdam, the Netherlands, 1963) pp. 236-243.

W.R. Ashby, "Induction, Prediction, and Decision-Making in Cybernetic Systems", in *Induction-Some Current Issues*, H.E. Kyburg and E. Nagel, eds. (Wesleyan University Press, Connecticut; AFOSRI 64-2006, AD 451576) pp. 55-56.

W.R. Ashby, *IEEE Trans. on Military Electronics*, "Systems and Information", MIL-7, pp. 94-97 (1963).

W.R. Ashby, *Cybernetica*, "Measuring the Internal Informational Exchange in a System", **1**, pp. 1-22 (1965).

W.R. Ashby and C.C. Walker, "Genius", in *Textbook in Abnormal Psychology*, P. London and D. Rosenhan, eds. (Holt, Rinehart, and Winston, Inc., New York) 1966.

W.R. Ashby, *Annual Review of Psychology*, "Mathematical Models and Computer Analysis of the Function of the Central Nervous System", **28**, pp. 89-106 (1966).

W.R. Ashby, "Modeling the Brain", in *Proceedings of the IBM Scientific Computing Symposium on Simulation Models and Gaming* (IBM, New York, 1966) pp. 195-208.

W.R. Ashby and C.C. Walker, *Kybernetik*, "On Temporal Characteristics of Behavior in Certain Complex Systems", **3**, pp. 100-108 (1966).

W.R. Ashby, *Currents Mod. Biol.*, "The Place of the Brain in the Natural World", **1**, pp. 95-105 (1967).

Dan Cohen, *J. Theort. Biology*, Optimizing Reproduction in a Randomly Varying Environment, **12**, pp. 119-129 (1966).

M.D. Freedman, *J. Aud. Eng. Soc.*, "A Digital Computer for the Electronic Music Studio", **15**, pp. 43-50 (1967).

M.D. Freedman, *J. Acoust. Soc. Am.*, Analysis of Musical Instrument Tones", **41**, pp. 793-806 (1967).

Gotthard Gunther and Heinz von Foerster, "The Logical Structure of Evolution and Emanation", in *Interdisciplinary Perspectives of Time*, R. Fisher, ed., (New York Academy of Sciences, New York, 1967) pp. 874-891.

A. Inselberg, *Notices Amer. Math. Soc.*, "Applications of Superpositions for Nonlinear Operators", **13**, number 3, pp. 357 (1966).

A.A. Mullin, *Nat. Acad. Sci.*, "Models of the Fundamental Theorem of Arithmetic", **50**, pp. 604-606 (1963).

Heinz von Foerster, "Logical Structure of Environment and its Internal Representation", in *International Design Conference, Aspen, Colorado, 1962*, R.E. Eckerstrom, ed. (Herman Miller, Inc., Zeeland, Michigan, 1963; AFOSRI 1542, AD 427409) pp. 27-38.

Heinz von Foerster, "Memory without Record", in *The Anatomy of Memory*, D.P. Kimble, ed., (Science and Behavior Books, Palo Alto, California, 1965; AFOSR 67-0311, AD 646844) pp. 388-433.

Heinz von Foerster, *Datamation*, "Computers in Music", **12**, number 10, pp. 106-111 (1966).

Heinz von Foerster, "How to Tell the Birds from the Bees: The Ontogenesis of Information", in Proceedings *IFIP Congress 1965*, W.A. Kalenich, ed., (Spartan Books, Washington, D.C., 2, 1966) p. 636.

Heinz von Foerster, "From Stimulus to Symbol", in *Sign, Image, Symbol*, G. Kepes, ed. (George Braziller, New York, 1966) pp. 42-61.

Heinz von Foerster, *Current Mod. Biology*, "Computation in Neural Nets", **1**, pp. 47-93 (1967; AFOSR 67-0982, AD 651673).

Heinz von Foerster, "Time and Memory", in *Interdisciplinary Perspectives of Time*, R. Fischer, ed. (New York Academy of Sciences, New York, 1967) pp. 863-873.

Paul Weston, *J. Data Management*, "Machine Use of 'Natural Language", **1**, number 4, pp. 32-35 (1963).

G.W. Zopf, "Sensory Homeostasis" in *Nerve, Brain, and Memory Models*, N. Weiner and J.P. Schade, eds. (Elsevier Press, Amsterdam, The Netherlands, 1963) pp. 114-121.

Doctoral Theses

M.L. Babcock, "Reorganization of Adaptive Automation" (Department of Electrical Engineering, University of Illinois, Urbana, 1960) 109 pp.

S. Cheng, "Life Span and Self-Repair in Complex Systems" (Department of Electrical Engineering, University of Illinois, Urbana, 1960) 53 pp.

M. Freedman, "A Technique for the Analysis of Musical Instrument Tones" (Department of Electrical Engineering, University of Illinois, Urbana, 1965) 138 pp.

J. Gazdag, "A Method of Decoding Speech" (Department of Electrical Engineering, University of Illinois, Urbana, 1966) 129 pp.

A. Inselberg, "Classification and Superposition Principles for Non-linear Operators" (Department of Mathematics, University of Illinois, Urbana, 1965) 71 pp.

H. Na, "On Structural Analysis of Many-Valued Logics" (Department of Electrical Engineering, University of Illinois, Urbana, 1964) 132 pp.

R. Swallow, "Cognitron Theory" (Department of Electrical Engineering, University of Illinois, Urbana, 1964) 243 pp.

I. B. Thomas, "The Significance of the Second Format in Speech Intelligibility (Department of Electrical Engineering, University of Illinois, Urbana, 1966) 156 pp.

C.C. Walker, "A Study of a Family of Complex Systems - An Approach to the Investigation of Organisms' Behavior" (Department of Psychology, University of Illinois, 1965) 251 pp.

Technical Reports

A. Andrew, "Table of the Stirling Numbers of the Second Kind" (Technical Report Number 6, 1965, 154 pp.; AFOSR 66-2373, AD 641081).

A. Andrew, "Automatic Adjustment in a Continuous Environment" (Technical Report Number 8, 1965, 19 pp.; AFOSR 65-2283, AD624548).

A. Andrew, "Modified Stirling Numbers of the Second Kind" (Technical Report Number 7, in preparation).

W.R. Ashby, "Constraint Analysis of Many-Dimensional Relations" (Technical Report Number 2, 1965, 17 pp.; AFOSRI 64-1387).

M.D. Freedman, "A Technique for the Analysis of Musical Instrument Tones" (Technical Report Number 6-13, 1965, 137 pp.; AFOSR 65-2293, AD 627181).

J. Gozdag, "A Method of Decoding Speech" (Technical Report Number 9, 1966, 19 pp.; AFOSR 66-2385,,AD 641132).

A. Inselberg, "On Classification and Superposition Principles for Nonlinear Operators" (Technical Report Number 4, 1965, 73 pp.; AFOSR 65-1712, AD 624268).

J. Russell, "The Visual Image Processor" (Technical Report Number 11, 1967, 121 pp.; AFOSR 67-1665, AD 655842).

H. Ryan, "Classification and Enumeration of Autonomous Sequential Machines" (Technical Report Number 12, 1967, 110 pp.; AFOSR 67-1412, AD 653999).

R. Swallow, "Cognitron Theory" (Technical Report Number 3, 1964, 243 pp.; AFOSR 65-0582, AD 614006).

I.B. Thomas, "The Significance of the Second Formant in Speech Intelligibility" (Technical Report Number 10, 1966, 155 pp.; AFOSR 67-1414 AD 654326).

C.C. Walker, "A Study of a Family of Complex Systems - An Approach to the Investigation of Organisms' Behavior" (Technical Report Number 5, 1965, 251 pp.; AFOSR 65-1713, AD 624864).

Laboratory Reports

A. Andrew, "Counting" (BCL Report, 1966).

W.R. Ashby and C.C. Walker, "Complex Systems (BCL Report 5.5, 1966, 7pp.).

BCL Staff, "Accomplishment Summary of the Biological Computer Laboratory for the Period 1, May 1965 - 31 April 1966 (BCL Report 66.1, 1966, 23 pp.).

BCL Staff, "BCL Publications" (BCL Report 67.1, 1967, 36 pp.).

BCL Staff, "Accomplishment Summary 66/67" (BCL Report 67.2, 1967, 126 pp.).

C.W. Burckardt, "From Flicker-Fusion to Color Vision" (BCL Report 5.3, 1966, 41 pp.).

D. Cohen, G. Brieske, H. von Foerster, "A Formal Model of Optimization in a Stationary Stochastic Environment" (BCL Report 7.1, 1967, 17 pp.).

G. Gunther, "Cybernetics and the Transition from Classical to Trans-classical logic" (BCL Report 3.0, 1964, 9 pp.).

G. Gunther, G.C. McVittie, H. von Foerster, "Time" (BCL Report 3.1, 1966, 50 pp.).

G. Gunther, "Formal Logic, Totality and the Super-Additive Principle (BCL Report 3.3, 1966, 23 pp.).

K. Kokjer, "Tables x $\log_2$ x for Itegers 1 x 1009" (BCL Report 1.0, 1966).

K. Kokjer, "Cutaneous Channel Capacity" (BCL Report 6.1, 1966, 17 pp.).

S. Powers, "Uncertainty Analysis in Dynamic Systems (BCL Report 8.0, 1967, 35 pp. ; AFOSR 67-1413, AD654310).

J. Russell, "A Visual Image Processor (VIP)" (BCL Report 5.0, 1965).

J. Russell, "Visual Image Processor II" (BCL Report 5.4, 1966, 65 pp.).

D.J. Schadoch, "A Classification of Mappings between Finite Sets and Some Applications" (BCL Report 2.2, 1967, 27 pp.).

Heinz von Foerster, "Neural Networks" (BCL Report 5.1, 1963, 126 pp.).

Heinz von Foerster, "From Stimulus to Symbol: The Economy of Biological Computation" (BCL Report 3.2, 1966).

Heinz von Foerster, "The Numbers of Man, Past and Future" (BCL Report 13.0, 1967, 22 pp.).

P. Weston and H. Ryan, "Rooted Trees" (BCL Report 2.0, 1966, 51 pp.).

THEORETICAL MODELS FOR THE RETICULAR FORMATION

William L. Kilmer

Michigan State University, East Lansing, Michigan

Project Description

The core of the central nervous system, called the reticular formation, functions as a highly sophisticated command and control center. Structurally, it consists of a string of similar modules, wide but shallow in computation, with connections not only from module to adjacent module, but between distant modules. The reticular formation receives relatively unrefined information from all sensory-motor systems and from all internal housekeeping systems. It processes crucial information across a broad front and arrives at a decision that commits the entire organism in a fraction of a second. This effort is developing a theory of its circuit action through increasingly advanced computer simulations of a mechanical reticular calculus of relations approximating the output modal behavior of the reticular formation. Strategies are being augmented to enable the model to habituate, condition, generalize, discriminate, trial-and-error learn, and follow a changing environment. Research is also planned on the postulated modality of vertebrate brain states indicated in EEG patterns that seeks to relate relevant aspects of the brain states to mathematically definable features.

Summary of Recent Research Results

A computer program has been written to test a trainable reticular formation model, and systematic testing will shortly begin. Emphasis will be on finding how much and how fast the model can learn, starting from zero knowledge of its environment. Proper response modes have been assigned for all allowable model inputs so as to match the assignment structure to the model's training methods. This yields a model environment in which two arbitrary model inputs that are close together in Hamming distance correspond to the same output mode with probability greater than 1/4, the pure chance level for 4 modes. This approximates animal-environment, stimulus-response relations for vertebrates fairly well.

This is not just another learning model investigation. It is a study of a cooperative decision-making system (of 12 modules) augmented to cooperatively organize its output mode behavior so as to cope with its environment.

97

Since reinforcement takes place after the response to each model input of reasonable duration, an important question is, can this model learn from scratch everything it could sustain if completely and optimally preprogrammed.

An attempt will be made to discover what the model can not learn with its training methods, and what the price is for having such simple reinforcement rules. Information will also be sought on how much and how fast the model can learn in an environment ideally mismatched to its training methods.

Reports and Publications

William L. Kilmer and J. Blum, "Some Mechanisms for a Theory of the Reticular Formation" (Final Report, February 1967; AFOSR 67-0928, AD651207).

98

4.10 CONFERENCE ON LEARNING, REMEMBERING, AND FORGETTING

Frank Fremont-Smith

New York Academy of Science, New York, New York

Contract Numbers* AF-49 (638) - 1770, F49604-67-C-154

Project Description

Results of fundamental research in the medical and behavioral sciences strongly suggest bases for a theory of teaching and for the understanding of learning problems. Research findings are seldom stated in terms that can be applied directly by teachers to classroom situations. Also, findings from one research project are often in terms and conceptual frameworks so different from the next that sophisticated specialists in one discipline have difficulty in interpreting results outside their field. This conference will provide a forum that will enable a selected group of specialists in the medical, behavioral, biological, and physical sciences to consider aspects of learning phenomena that transcend their disciplines. The conference is the fourth in a series of five. Published proceedings will make the communications and interchanges among the specialists readily available to advance the state of knowledge and understanding of learning problems.

Summary of Recent Research Results

The fourth annual interdisciplinary Conference on Learning, Remembering and Forgetting (sponsored jointly by AFOSR and the New York Academy of Sciences Interdisciplinary Communications Program) was held October 9-12, 1966 in Pacific Palisades, California. The conference was divided into three main sessions. The main topic of discussion of the sessions was the development and differentiation of behavioral capacity. This major topic was discussed from several empirical view points that ranged from experiments involving perceptual and perceptual-motor rearrangement to the development of species-specific behaviors, and genetic and hormonal bases for learning and memory capabilities.

Drs. Richard Held and Alan Hein initiated the first day's discussion by presenting their recent work on the importance of certain forms of contact with the environment for the development of

normal spatially oriented responses. Dr. Horst Mittelstaedt of the Max-Planck Institute contributed complementary ideas regarding the important role of corallary neural discharge for coordinated visually guided behavior.

Distinctions among various modes of action during early life on both sensory-motor and perceptual discrimination were proposed by Dr. Robert A. Hinde as the base for the second day's discussion. The discussion evolved to include a detailed consideration of the growth of linguistic characteristics in human infants by Dr. Alvin M. Liberman.

Drs. Seymours Levine and Robert W. Goy oriented the final day's session by discussing the differing roles of hormones on the development of the organism. The data presented seemed to indicate that during certain sensitive periods the central nervous system is subject to the organizing effects of these endocrines. This is revealed by later sexual differentiation in morphology and more surprisingly, in the "maleness" or "femaleness" of the animal's behavior.

The proceedings of the conference are presently being edited by Dr. Daniel P. Kimble.

100

4.11   SELF-EXTENDING COMPUTER LANGUAGES AND THEIR TRANSLATION

Richard K. Bennett

Signatron, Inc., Lexington, Massachusetts

Project Description

The manner in which computer programming languages have developed makes it impractical for a computer user to modify or alter the built-in facilities of the particular language he is using. Thus, for example, a user of the FORTRAN language is essentially restricted to writing statements and using data structures which are defined in FORTRAN. He cannot modify the language in any way to provide new instructions and data representations that are directed to his immediate problem. The proposed research will attempt to create a new direction in the development of programming languages, and is based on successful initial research results.

In this approach, the user is provided with a language base on which he may build a language tailored to his own (or his group's) requirements. Or the user may instead start from an existing language built on this base and extend or modify its facilities.

This language base consists of a set of ground rules adequate to base a class of languages broad enough to cover essentially all computer applications. In addition, the base includes all of the facilities and devices which have proved useful in existing languages, such as algebraic expression handling and list processing. A language is defined on this base simply by defining the set of verbs of the language, stating what function each is to serve and what modifiers (parameters) each may take. The statement structures in this approach are greatly simplified, so that the user is normally unconcerned with syntactic considerations. (Complex syntactic forms can be handled as special cases, if the user so desires.)

Thus, it should be a simple matter to define, say, a civil engineering language. Instead of requiring a large-scale effort as it does today, the user would simply define the functions he desires, such as laying out lines and arcs, computing areas, constructing tangents, etc. Since the user would be drawing on the very powerful and proven facilities of the language base (and its translator), he would not have to re-create such operations as reading character strings, searching symbol tables, and evaluating expressions. His job then would be one of weeks or of months, instead of years.

101

This approach potentially has an even larger payoff. The languages built on this language base would all be *compatible*--could all be translated by the same translator. Thus there exists the possibility of eliminating, or at least reducing, the costly "Tower of Babel" problem which has grown to such great proportions during the last few years.

Summary of Recent Research Results

This AFOSR research effort received its initial sponsorship less than one year ago, and will be reported in the next issue of this publication.

Reports and Publications

Richard K. Bennett, "A Base for the Definition of Computer Languages" (Signatron Technical Report, October 1967).

4.12 CYBERNETIC INVESTIGATION OF LEARNING AND PERCEPTION

Gordon Pask

System Research Ltd., Richmond, England

Project Description

This effort is exploring mechanisms, principles, and algorithms by which learning can be described, as mediated by the human brain, biological structures, and machines. A learning or perceptual system is conceived as a self-organizing system, the ordering relations being derived from interaction between the learning or perceiving organism or mechanism and its environment. The essential parts of a model for concept learning have been defined as a hierarchy of problem-solving processes, a hierarchy of mechanisms corresponding to the process hierarchy, and levels of metalanguage with reference to an object language. Models have been simulated for populations of automata that become more complex as they develop. Stability was shown over a larger range of cost parameter values for the population in an unconstrained in contrast with a constrained environment. Toward specifying fundamentals of the learning process, research is continuing on simulations of the acquisition of simple subskills, the use of unstratified languages that minimize pragmatic ambiguities as learning occurs over a hierarchy of levels of discourse, the behavior of populations of automata postulated to function as cooperative social systems, and the role of sequential substitution and chain learning in cybernetic models.

Summary of Recent Research Results

During the past year our basic simulation program has been re-organized. In its present form, it is capable of modelling various types of boredom and fatigue that are observed in the acquisition of code application and code alternation skills (the latter situation involves a mechanism for changing the "attention" of the model). Although the broad features of the model can be (and have been) checked against experiments with real subjects, there are certain detailed modes of behavior that call for special experimentation. The equipment required for this purpose has been built, and the experiments are being conducted.

The basic model has also been modified to account for the acquisition of skills such as teleprinting in which the subject assembles a sequence of operations to produce a response.

103

Since this type of learning is closely related to verbal learning, we have posed a number of hypotheses and are currently attempting to bridge the gap between perceptual motor and verbal behaviour. The theoretical consideration that underlies this effort is the contention that the stimulus response paradigm is inadequate to account for even simple skills; many aspects of human learning appear to involve programmatic components (control units assembled into a cognitive structure) that are embodied in the innately structured computing machine of the br.in.

The question of analogy operations (mooted in previous theoretical reports) has appeared in the course of the simulation research in connection with stimulus equivalence and the equivalence between sequences of problem solving operations produced by the system. We are currently dealing with this question by: (1) adding additional parts to the executive routines of the model; and (2) Incorporating an evolutionary process which generates (and, in conjunction with the existing model, selects) stimulus descriptions and operations that are analogous to those already in the model's repertoire.

These processes involve a primitive sort of innovation. Next year we hope to simulate less primitive types of innovation and to perform experiments (with human subjects) in which some of the sub-processes can be externalized and observed.

Reports and Publications

Gordon Pask, "Statistical Computation and Statistical Automata", Second Conference on Problems in Cybernetics, Karlsruhe, Germany (1-18 April 1963; AFOSRJ 586).

Gordon Pask, B.N. Lewis and D. Watts, "A Typical Adaptively Controlled Experiment in Perceptual Discrimination" (Supplementary Technical Report, November 1964; AFOSR 65-0291, AD611542).

Gordon Pask, "Research on Cybernetic Investigation of Learning and Perception" (Supplementary Technical Report Number 2, December 1964; AFOSR 65-0292, AD611543).

Gordon Pask, "A Discussion of Artifical Intelligence and Self-Organization" in *Advances in Computers*, F.L. Alt and M. Rubinoff, Eds., 5, (Academic Press, New York, 1964; AFOSR 65-1168, AD620496) pp 110-225.

Gordon Pask, *Control Automation Progr.*, "Teaching as a Control-Engineering Process, Part 1", 9, Number 79, pp. 6-11 (January, 1965; AFOSR 65-0583, AD618003).

Gordon Pask, "Some Comments on the Organization of Men, Machines, and Concepts", in *Proc. Am. Doc. Inst. Working Symp. Educ. Inform. Sciences* (at Airlie House, Va., 7-10 September 1965; AFOSR 65-2666, AD628141) pp 133-154.

Gordon Pask, *Bulletin of Mathematical Biophysics*, "Man-Machine Interaction in Adaptively Controlled Experimental Conditions", 27, (1965, special issue).

Gordon Pask, M. Elstob and G.L. Mallen, "Research on Cybernetic Investigation of Learning and Perception" (Annual Summary Report Number 3, February 1966; AFOSR 66-0644, AD631634).

Gordon Pask, Philosophical Discussion of Cybernetics at the Technical University of Berlin (Cybernetics Seminar, July 1966; to be published).

Gordon Pask, "A Cybernetic Model for Some Types of Learning and Mentation", to be published in *Cybernetic Problems in Bionics,* Gwinn, Gordon and Breech, Eds. (presented at Bionics Symposium, Dayton Ohio, 1966).

Gordon Pask, "Intelligent Learning", in *Proceedings of the Bionics Symposium* (Dayton Ohio, 1966).

Gordon Pask, *Progress in Biocybernetics,* "Comments on the Cybernetics of Ethical, Sociological and Psychological Systems", 3, (Elsevier Publishing Company, Amsterdam, 1966).

Gordon Pask and R.J. Feldmann, *Cybernetica,* "Tests for a Simple Learning and Perceiving Artifact", 2, (1966).

Gordon Pask and G.L. Mallen, "The Method of Adaptively Controlled Psychological Learning Experiments", in *Theory of Self-Adaptive Control Systems* (Plenum Press, 1966).

Gordon Pask, "Research on Cybernetic Investigation of Learning and Perception" (Final Report, January 1967; AFOSR 67-0861, AD650290).

Gordon Pask, "Experiments in the Simulation of Learning and Mentation", in *Proc. 5th Intl. Cong. of Cybernetics* (Namur, 1967).

Gordon Pask, "Recent Work with Hierarchically Organized Teaching Machines", in *Proc. 5th Intl. Cong. of Cybernetics* (Namur, 1967).

Gordon Pask, "Adaptive Teaching Systems", British Association for the Advancement of Science (University of Leeds, England, 1967).

Gordon Pask and B.N. Lewis, "The Adaptively Controlled Instruction of a Transformation Skill", in *Programmed Learning,* Sweet and Maxwell, Eds. (April 1967) pp. 74-86.

Gordon Pask, *La Cybernetique,* "Some Difficulties Encountered in Psychological Experiments on Learning", Number 26 (June 1967).

Gordon Pask, "The Control of Learning in Small Subsytems of a Programmed Educational System", in *IEEE Trans. of Human Factors in Electronics,* H.F.E.8 (June 1967, Special Issue).

Gordon Pask, "A Learning Model Capable of 'Attention' and Hampered by 'Boredom' and 'Fatigue' ", in the *NATO Symposium on the Simulation of Human Behavior* (Paris, July 1967).

Gordon Pask and B.N. Lewis, "The Self-Organization of a Three-Person Task Oriented Group", in the *NATO Symposium on the Simulation of Human Behavior* (Paris, July 1967, to be published in proceedings).

Gordon Pask, "Adaptive Teaching Machines," (Lecture and Summary delivered at the UNESCO Seminar, Turin, September 1967).

Gordon Pask and B.N. Lewis, "British *Journal of Mathematical and Statistical Psychology,* "Studies in the Acquisition of Simple and Complex Transformation Skills Using a Null Point of Steady State Method", (to be published).

Gordon Pask, "Cybernetics and Education" (paper to be given during the 10th Anniversary of the Intl. Assoc. for Cybernetics).

Gordon Pask, G.L. Mallen and M. Robinson, "Some Difficulties and Inconsistencies Encountered in Connection with Adaptively Controlled Teaching Systems" (in preparation for *Programmed Learning).*

Gordon Pask, *Scientia,* "The Relation Between Behavioristic and Cybernetic Approaches to Psychology" (to be published).

Gordon Pask, A Course of 10 Lectures on "Teaching Systems" and 10 Lectures on "Behavoral Cybernetics" given at the Institute of Advanced Studies, Department of Engineering, National University of Mexico.

106

4.13   **MACHINE LEARNING FOR
GENERAL PROBLEM SOLVING**

David L. Johnson

University of Washington,
Seattle, Washington

Project Description

The ability to perform complex, repetitive, and recursive operations
rapidly and accurately offers man the possibility of utilizing the
computer in complex problem-solving and decision-making tasks. The
degree to which machines can be made useful depends on man's
ability to describe problems in machine-understandable terms and
his ability to describe machine models capable of problem solution.
This effort investigates the formulation and implementation of
computer-based methodology for several specific problem-solving
tasks. General learning processes and modeling functions may be
discovered by evaluating learning techniques within specific problem
areas. Coordinated investigations will be conducted into such areas as
pattern recognition, theorem-proving, and game learning. The possi-
bility of developing self-improving heuristics as a learning tool will
be investigated (i.e., a learning capability within a learning scheme).
Concept formation within child development is also being studied for
possible methods of computer modeling. The goals of this effort are
the formulation of optimum problem-solving routines and the under-
standing of man-machine interface areas in various tasks.

Summary of Recent Research Results

A tree pruning system (TPS) has been developed that consists of a
set of system statements to be used as instructions for programming
of heuristic tree searches. The system is imbedded in FORTRAN,
designed to treat non-uniform tree structures. Search decisions are
automatic, either on the basis of user overall specification or machine-
learned criteria. The basic application leading to the development of
the TPS is chess. System value is demonstrated by successful and
simple duplication of the Baylor and Simon "Chess Mating Combi-
nation Program".

107

An adaptive pattern classifier applicable to curve fitting problems has also been developed. The research contains additional theory for general pattern classifiers; i.e., one-step correction under various training methods, and development of a modified Euclidean metric. Two adaptive pattern classifiers have been designed for solutions of practical scientific problems. Heuristics for learning are coupled to threshold logical theory. In application the curve fitting model classifies realistic patterns at 200 to 300 patterns per minute on the Burroughs B5500 with average error of 3.5%. The new method does not require the "good" initial curve coefficient guesses required by other methods.

Investigations continue in various aspects of concept learning and formation. Only a minimal effort has been directed toward study of the man-computer relationship in respect to the effect of attitudes upon the man-computer interface.

Reports and Publications

Alistair D.C. Holden and David L. Johnson, "Simulation of the Growth of Scientific Reasoning in the Child", in *Proceedings of the Bionics Symp.*, Dayton, Ohio (1966).

Allistair D.C. Holden and David L. Johnson "The Use of Imbedded Patterns and Canonical Forms in a Self-Improving Problem Solver", in *Proceedings of the 22nd National ACM Conference*, pp. 211-219 (1967).

David L. Johnson and A. L. Kobler, *Science*, "The Man-Computer Relationship", **138**, No. 3543 (Nov, 1962).

David L. Johnson and A. L. Kobler, "Man-Computer Interface Study" (Final Report, June 1963; AFOSR 1540).

David L. Johnson, "Machine Learning for General Problem Solving". (Final Report, Oct 1964; AFOSR 64-2290).

David L. Johnson, Alistair D.C. Holden, E. W. Kozorowicki, G. Syms, and G. Moffitt, "Machine Learning for General Problem Solving", (Final Report, Apr 1966; AFOSR 66-0835, AD632576).

David L. Johnson and Alistair D.C. Holden, "Computer Learning in Theorem Proving", in *IEEE International Convention Record*, Part 6 (1966).

E. Kozdrowicki, "An Adaptive Tree Pruning System: A language for Programming Heuristic Tree Searches" (Technical Report; in press).

T. A. Marsland, "The Synthesis of Threshold Logic Networks through the Evolution of an Adaptive Computing System" (Technical Report; in press).

G. H. Syms, "A Pattern Recognition Model for On-Line Curve Fitting: An application of Threshold Theory" (Technical Report; in press).

Section 5

LANGUAGE AND LINGUISTICS RESEARCH

Language consists of groups of symbols arranged according to a set of rules. Viewed this way, "language" includes both the languages people speak and all synthetic symbol systems. So language is viewed in this task.

Language is a tool. It is a device for representation. Information science is concerned with its use for representing concepts, relationships among concepts, models for automata, and instructions for processing devices.

Natural language has proved tractable with difficulty. Structure or syntax has been amenable to modeling, but the number of grammatically correct sentences that can be meaningless is for practical purposes, infinite. Major emphasis is now on semantics, to elicit rules for recognizing meaning and generating meaningful sentences. Research on natural language was originally sponsored (not by AFOSR) to develop procedures for machine translation. Machine production of translations, indexes, and abstracts, and other machine manipulations of natural language text are requirements if achievable, because man's capacity to assimilate bits of information is limited. As document volume increases, automatic methods that intelligently select the significant bits appear mandatory for efficient handling of information. Results are slow in coming because the data is large, largescale projects are expensive, and perhaps creative ideas are few.

Work on synthetic or artificial languages can be roughly subdivided into work on machine and programming languages and work on languages for representing concepts and procedures simulating intelligent processes. The directorate program has not included much programing language research. An area of directorate interest is in translators and compilers which make machines accessible to non-programmer scientist and manager users.

Languages for game playing theorem proving suggest approaches to the representation and association of concepts that may be approximations to characterizing thought processes. Considerable progress has been made, some under AFOSR sponsorship, on languages for problem-solving procedures. Enough may now be known about some languages for automata to permit generalizations on useful procedures that can be accomplished on computers.

109

5.1 RESEARCH IN SECOND-LANGUAGE LEARNING

John A. Swets

Bolt Beranek and Newman Inc., Cambridge, Massachusetts

Current Contract Number* F44620-67-C-0033

Project Description

Learning a language requires learning rules pertaining to sounds
and their combination, rules of syntax, and rules of reference or
meaning. A second language introduces new acoustic and physiolog-
ical dimensions and other difficulties with phonetics and processes
of perception. This effort will analyze the processes of learning the
rules of a second language to aid the design of teaching procedures
and to determine the feasibility of developing an advanced computer-
based system to aid in language instruction. Attention will focus on
phonetics, primarily, and syntax. A system is envisioned that will
enlarge language-laboratory training by a system that performs an
automatic acoustic analysis of vocal, productions, evaluates the
correspondence between standard utterances and students' efforts,
displays the evaluation to the student, and enables presentation of
succeeding items in the lesson in accordance with the student's
performance. Research should enable the inventorying of acoustic
distinctions in various languages, a determination of those most
amenable to machine analysis, and a determination of effective
modes of feedback to students in man-machine situations.

Summary of Recent Research Results

In operation now is a computer-based system that enables a
researcher to define a variety of properties of the speech waveform
in terms of the outputs of a bank of filters. Guided by an inventory,
and some theory, of the acoustic distinctions difficult for a student
to learn in going from one particular language to another, and using
a list-processing computer language which facilitates program
modifications, data are being collected to evaluate the contribution
to analysis made by each of many specifiable properties of the
waveform. Oscilloscopic displays to indicate to the student the
nature of his errors in pronunciation are also functioning and in the
process of evaluation. A second computer-program system function-
ing now in a preliminary form is devoted to teaching syntax and

semantics. This system administers several modes of practice and testing, assesses critical aspects of student performance, and presents material conditional upon past performance. In final form, this system should be able to accept typewritten inputs rather freely generated by the student, detect any syntactic or semantic errors, and recognize many of the errors. A survey report of language-teaching methods has been prepared which highlights the influence of descriptive linguistics and the implications of the new generative-grammar theory.

Reports and Publications

John A. Swets, "Research in Second-Language Learning" (BBN Report No. 1530, Semiannual Technical Report No. 1, June 30, 1967).

5.2   1966 LINGUISTIC INSTITUTE
CONFERENCE ON LINGUISTIC METHOD

Paul L. Garvin

Bunker-Ramo Corporation,
Canoga Park, California

Project Description

The field of linguistics is becoming increasingly significant to
Information Sciences because of the need to apply both theoret-
ical linguistic concepts and methodological principles to the
analysis natural languages and the synthesis of artificial languages
processing by machines. To assess the present stage of development
of linguistics as an information science, this effort sponsors a
conference that will consider: (a) the relation between method
and theory in current linguistics research, (b) the methodological
procedures and techniques being used for various information
processing purposes, and (c) applications of empirical method to
problem areas in linguistics. Invited participants include exponents
of various approaches to theory and method in linguistics. The
conference is expected to contribute to a clarification of funda-
mental problems concerned with the scientific treatment of
natural language data and the development of formalizations of
both theoretical concepts and methodological procedures.

Summary of Recent Research Results

This conference was held August 1-3, 1966 at the University of
California, Los Angeles in conjunction with UCLA's annual
Linguistic Institute. Approximately 225 persons were in attendance.
The conference was organized into three general sessions, each
session occupying one day. The sessions were entitled "The Re-
lation of Method to Theory", the "Development of Methodological
Principles", and the "Applications of Linguistic Method." Indiana
University Press will publish the proceedings of the conference
under the title, "The Place of Method in Linguistics." The pro-
ceedings will contain an evaluation of the conference by Dr. Paul
Garvin, its organizer.

5.3 STUDIES IN MATHEMATICAL AND COMPUTATIONAL LINGUISTIC RESEARCH

Harold P. Edmundson

University of California, Los Angeles, California

Project Description

High-speed computers are being increasingly viewed by linguists as a useful tool for the analysis of large corpora of natural language and the complex interrelationships of the data uncovered. Concomitantly, mathematical models are being formulated and methods applied in the study of the language data. This effort examines the fundamental concepts of mathematical and computational linguistics and the relationships of these concepts to the various problems in language data processing. Theories founded in mathematical logic, analysis, algebra, geometry, probability, and statistics are being investigated and contrasted. A study is planned of computer routines and programs, including general-purpose routines for search, sort, lookups, and frequency; special routines for parsing, segmentation, and idioms; and programs for dictionary, thesaurus, and concordance compilation. Also, the complementary roles of theoretical models of linguistic phenomena and empirical language data are being examined. Results are expected to contribute ultimately to a more effective use of computers in such language data processing research as automatic indexing, abstracting, and translating.

Summary of Recent Research Results

Recent research results lie in both mathematical and computational linguistics. In mathematical linguistics, a sematic study has resulted in a new axiomatic characterization of synonymy and antonymy. The relations of synonymy and antonymy were regarded as ternary relations and their domains and ranges were examined. Synonymy and antonymy were defined jointly and explicitly by a set of eight axioms rather than separately by three axioms each as in an earlier study on this project. This axiom system has permitted the proofs of several theorems that satisfy linguistic intuition. In addition, two topological characterizations were posed and examined.

In computational linguistics a description of previous research on a system of automatic extracting was reported. The goal of this research was to develop a computer program for producing indicative, rather than informative, abstracts for the screening of scientific

113

documents for Air Force use. Although the initial system was programmed for an IBM 7094, it has been reprogrammed also for the Univac 1103A and CDC 1600. Moreover, the system for extracting English has been suitably modified to extract Russian scientific text. The final report contains the details of both the operating system and the research methodology.

Reports and Publications

Harold P. Edmundson, "Mathematical Models of Synonymy" in *1965 Intern. Conf. Computational Linguistics* (New York, N.Y., May, 1965; AFOSR 65-1387, AD621160).

Harold P. Edmundson, "New Methods in Automatic Extracting" (Report Number TM-3207, Systems Development Corporation, 1966).

Harold P. Edmundson, "Axiomatic Characterization of Synonymy and Antonymy" in *1967 Intern. Conf. Computational Linguistics* (Grenoble, France, August 1967).

Harold P. Edmundson, "Mathematical Models in Linguistics and Language" in *Automated Language Processing: The State of the Art,* H. Borko, Ed. (Wiley, New York, N.Y., 1967).

Harold P. Edmundson, "Mathematical and Computational Linguistics" in *Conceptual Bases and Applications in the Communication Sciences,* E. Beckenbach, Ed. (McGraw Hill, New York, to be published).

Invited Talks

"Automata and Their Associated Languages" Mathematics Symposium, Harvey Mudd College, Claremont, California, Sept. 1966.

"Language Data Processing", Amer. Doc. Inst. Convention, Santa Monica, Calif., Oct. 3, 1966.

"Linguistics and Artificial Intelligence", SIGART Meeting, Assoc. for Computing Machinery Los Angeles, Dec. 1966.

"Relations between Computers, Language, and Statistics", Amer. Stat. Assoc. and Assoc. for Computing Machinery Symposium on Computers and Statistics, Santa Monica, Calif., Feb. 1967.

"Natural Language Processing and the Library" Information Science and Librarianship Symposia, Univ. of Wisconsin, March 1967.

5.4   **PROBLEM SOLVING AND PATTERN RECONGNITION LANGUAGES**

Ranan B. Banerji and George Ernst

Case Institute of Technology,
Cleveland, Ohio

Project Description

Machine recognition of situations as belonging to particular classes or as imposing a preferential ordering among possible transformations presupposes that the machine has a language that permits description of the situation. This yields the requisite problem-solving or decision making capability. Prior research has produced a spectrum of languages with increasing flexibility. These have been able to represent problems and the systemizations of experience. A subclass of context-free languages (the finitely representable) has been isolated. A context-free language has been developed capable of describing sets, relations and operations with equal ease in terms amenable to machine manipulations. The methods are directed towards enabling a machine to discover properties of situations in terms of which strategic concepts have simple descriptions; large permentation groups under which a strategic concept is invariant; discovery of useful transformation through an analysis of their representations; similarities and differences between situations and the relevance of transformations to reducing particular differences.

Summary of Recent Research Results

A method has been developed and applied for learning strategic sets over a wide class of board games of the "Hex" variety. The success of the method is basically due to the use of the language of description. Well-known problem solving methods have been given mathematical form. The application of Automata and Graph Theory has yielded a method for decomposing games to yield efficient winning strategies.

Reports and Publications

E.B. Altman, "The Concept of Finite Representability"
(Supplementary Tech. Report No. SRC 53-A-64-19, November
1964; AFOSR 64-2408, AD609420).

115

Ranan B. Banerji, "Computer Programs for Generation of New
Concepts from Old Ones" (Symposium on "New Developments
in Cybernetics", Karlsruhe 1963, Oldenbourg Verlag).

Ranan B. Banerji, *Information and Control*, "Phrase Structure
Languages, Finite Machines and Channel Capacity", **6**, p. 153
(1963).

Ranan B. Banerji, "A Language for the Description of Concepts",
in *Yearbook of the Society for General Systems Research*, **9**,
p. 135 (1964).

Ranan B. Banerji, "Toward a Formal Language for Describing
Object Classes", in *Proc. 2ND. Congr. Inform. Sys. Sci.*, 22-25
Nov 1964, Washington, D.C. (Spartan Books, Inc., 1965;
AFOSR 66-0420, AD632774) pp. 451-457.

Ranan B. Banerji and T.G. Windekrecht, *Zeitschrift Fur.Psychologie*,
"Neure Ergebrisse Zur Theorie der Begriffsbildung", **171**, p. 204
(1965, Leipzig),

Ranan B. Banerji and E.B. Altman, *Information and Control*,
"Problems of Finite Representability", **8**, p. 251 (1965).

Ranan B. Banerji, "A Formal Model for Concept Description and
Manipulation" (Symposium 18 - The Organization of Human
Information Processing - in the 18th International Congress
of Psychology, Moscow, 1966).

Ranan B. Banerji, "Relation of Grammar to Efficiency in a
Parsing Algorithm", in *Computation in Linguistics - A Case
Book*, Garvin and Spolsky, Eds. (Indiana University Press,
1966) p. 76.

Ranan B. Banerji, *Journal of the Pattern Recognition Society*,
"A Language for Pattern Recognition" (to be published).

R.L. Citrenbaum, "The Concept of Strategy and its Application
to Three - Dimensional Tic-Tac-Toe" (Supplementary
Technical Reports No. SRC 72-A-65-26).

G.J. Hansen, "A Heuristic Approach to Solutions of Checkerboard
Puzzles (Supplementary Technical Report No. SRC 49-A-64-16,
November 1964; AFOSR 64-2406, AD609570).

M. Hueckel, "A Learning Scheme as a Possible Basis for Artificial
Intelligence" (Supplementary Technical Report No. SRC
94-A-66-38).

E.G. Koffman, "Learning Through Pattern Recognition Applied
to a Class of Games" (Supplementary Technical Report No.
SRC 107-A-67-45).

E.G. Koffman, "Learning Games Through Pattern Recognition",
in *Proceedings of the First Annual Princeton Conference on
Information Sciences and Systems*, p. 104 (1967).

L.R. Marino, "Winning and Non-Losing Strategies in Games and
Control" (Supplementary Technical Report No. SRC
91-A-66-36).

T.J. Mazuchowski, "Towards a Formal Theory of Language
Translation" (Supplementary Technical Report No. SRC
97-A-66-40).

C.E. Milliken, "A Computer Program for a Description Language"
(Supplementary Technical Report No. SRC 92-A-66-35).

116

J. Pennypacker, "An Elementary Information Processer for Object
  Recognition" (Supplementary Technical Report, No. SRC
  30-I-63-1).
J. Snediker, "A Generalized Program for Information Retrieval"
  (Supplementary Technical Report No. SRC 97-A-66-40).
T.G. Windeknecht, "A Theory of Simple Concepts with
  Applications"(Supplementary Technical Report No. SRC
  53-A-64-19; AFOSR 64-2407, AD609421).

Invited Lectures

Ranan B. Banerji, "A computer language for concept description"
  at David Sarnoff Research Laboratory, R.C.A.; Princeton, N.J.
  (1965).
Ranan B. Banerji, "A pattern recognition language", Computer
  Science Seminar, at The Pennsylvania State Unversity (1965).
Ranan B. Banerji, "Learning Games by Pattern Recognition"
  Computer Science Seminar, at University of Tornonto
  (February, 1967).
Ranan B. Banerji, "A Systems Approach to Two-person Games"
  Symposium on Mathematical Systems Theory at the Pennsylvania
  State University (September, 1967).
Ranan B. Banerji, "Solution of Two Person Games by Decompo-
  sition" Seminar on Computer and Their Capabilities at the Moore
  School of Electrical Engineering, The University of Pennsylvania,
  Philadelphia (October 1967).

5.5 SURVEY OF ON-LINE COMPUTER LANGUAGES AND SYSTEMS

Lewis C. Clapp

Computer Research Corporation, Newton, Massachusetts

Project Description

Languages and procedures by which problems can be described and represented for machine processing have proliferated in recent years. Between 400 and 500 languages, language variants, and schema may now be in use or development. Languages and systems are designed to satisfy problem solving needs, but too often they are not publicized and remain unknown to most of the growing community of machine users. This effort will provide a comprehensive survey and analysis of existing on-line computer languages and systems. Data will be obtained on their application areas and functional descriptions, the host computer system, the implementation language, input/output device specifications, literature references, numeric, dimensions of the systems, and the locations where the languages and systems are being used. Trends and patterns discovered during data classification and analysis will be noted. Results are expected to benefit present and potential users of machines and researchers and developers of new on-line systems.

Summary of Recent Research Results

As a first step toward obtaining information for the report, questionaires were sent out to 565 organizations working in a variety of disciplines and using on-line systems. Three hundred and seventy-five replies have been received to date. This has been supplemented with information obtained by visits to major research centers. A report breaking down the information into major functional categories; such as on-line computation, on-line graphics etc. is now in the final stages of preparation. The information about each system or language has been cross indexed by major system characteristics so that the reader can quickly locate all systems which use a particular technique, computer, language or terminal device.

To do justice to the large amount of detailed technical material collected, which could not be covered in a single report, we have started work on a second report which will review current techniques of on-line system development.

Reports and Publications

Lewis C. Clapp, et al, "A Survey of On-Line Computer Languages and Systems" (Report in preparation for Contract F44620-68-C0015).

Lewis C. Clapp, "Techniques in On-Line System Development" (Report in Preparation for Contract F44620-68-C0015).

119

5.6 A STUDY OF CERTAIN PHILOSOPHICAL INTRUSIONS IN CURRENT LINGUISTIC THEORY

Alice Koller

Connecticut College, New London, Connecticut

Project Description

This research takes as its starting-point the hypothesis that linguists have made implicit use of certain philosophical positions in developing their theories and hypotheses about language. The research project is an attempt to set forth the philosophical issues involved by closely examining the work of linguistic theorists. Such an examination may persuade linguists of the extent to which they have unknowingly taken philosophical positions on philosophical issues. Linguists will then be more easily able to rid linguistic theory of some of the assumptions which currently inhibit its progress.

Summary of Recent Research Results

Some beliefs about language and linguistics were critically analyzed. The beliefs, to which most contemporary linguists assent but which nevertheless hinder their theoretical work, are: that the data of linguistics are concrete physical objects; that linguistic method is primarily descriptive; that linguistic laws are laws of nature comparable to physical laws; that linguistics is an empirical science using inductive methods; that (contrariwise) linguistics is a formal science comparable to logic; that languages are codes; that the notion of correctness has no place in linguistics (particularly, that linguistics provides no basis for statements about correct usage); that the meanings of words are arbitrary; that meanings are associations and hence subjective; that (contrariwise) meanings are things-out-in-the-world and hence too plentiful; that individual words can be either ambiguous or meaningless; that words point to reality; that only nouns, verbs, adjectives and adverbs are meaningful; that linguists need to invent their own theory of meaning before they can get on with their work; that (contrariwise) linguists need have nothing to do with the meanings of words; that the notion of units of meaning both makes sense and is useful; that all sentences possess information content; that the notion of well-formedness adds something important to grammatical theory; that structure takes priority over meaning in constructing a grammar; and that speaking a language depends upon the prior reconstruction of the theory of that language.

120

The purpose of the analysis was to invite linguists to free themselves from these beliefs which rest ultimately upon certain philosophical commitments which linguists have accepted without assessing them properly.

Reports and Publications

Alice Koller, "A Hornbook of Hazards for Linguists" (April, 1967, AFOSR 67-0770).

Ernst Von Glasersfeld

Georgia Institute for Research, Inc.,
Athens, Georgia

Project Description

An ultimate goal of research on natural languages is to obtain
programs or sets of instructions by which machines will be able to
automatically process natural-language texts that constitute the
information base for various communication, documentation, and
management systems. Machines with such capability could then
perform such operations as translation, indexing and abstracting
automatically. Prior research on this effort has resulted in a
natural language analysis procedure that identifies syntactic and
semantic factors by means of a predictive correlational grammar.
The meaning of individual words and the relations that link them in
a sentence are represented by a hierarchic structure of correlations.
Research is continuing on the syntactic and semantic characterization
of words that will enable their significant concatenation in sentences.
Implicitly stated relationships and a (supralinguistic) classification
of correlators based on a comparative analysis of natural languages
will be examined. The procedure is being implemented on an IBM
360/65. Computer outputs will be analyzed for factors requiring
disambiguation. A weighting system will be studied for the
further resolution of ambiguities.

Summary of Recent Research Results

The Multistore procedure (an automatic parsing program for
English sentences) a first version of which (MP1) had been
implemented on a GE 425 computer, has been revised, expanded
and re-written for an IBM 360/65. The new program, MP2, is
at present being debugged and will be fully operational before long.
Among the advances incorporated in MP2, three would seem to be
particularly significant: 1) Extension of the 'significant address'
method to all matching-operations in the analysis procedure -
which means that there is no longer any need to shift data in the
machine, after input; 2) The assignation of predictive indices to
the word combinations produced in the course of an analysis is
now achieved by a homogeneous system of binary screens, which is
extremely economical and, consequently, fast; 3) A new type of

---

rule which prevents the formation of word combinations which, although both syntactically and semantically acceptable, are necessarily superseded in the particular sentence under analysis.

Reports and Publications

Brian Dutton and Ernst v. Glasersfeld, "Supralinguistic Classification and Correlators" in *Beiträge zur Sprachkunde und Informationsverarbeitung* (November, 1966).

Brian Dutton, "Some Causative Verbs in English and Arabic, an Initial Approach to a Semantic System", in *Final Scientific Report* (Contract F44620-67-C-0028, April 1967).

Ernst von Glasersfeld, "Automatic English Sentence Analysis" No. 2, Part 1 (Final Report, Grant AF-EOAR-65-76, September, 1966).

Ernst von Glasersfeld, Final Scientific Report, Contract F44620-67-C-0028 (April, 1967).

Ernst von Glasersfeld and Pier Paolo Pisani, "The Multistore Parser", Monograph to be published in the spring of 1968.

Ernst von Glasersfeld, "A Flexible Representation for the Meaning of Activity-Words in Natural Languages" (in Preparation).

Brunella Notarmarco, "Correlational Characteristics of English Adverbs of Degree" (in Preparation).

5.8 CRITERIA FOR EXCLUSIONS AND CORRELATIONS FOR AUTOMATIC LANGUAGE ANALYSIS

Jehane B. Burns

Girton College, Cambridge, England

Project Description

Machine programs for the manipulation of natural language text
for such purposes as automatic indexing, abstracting, or translation
will be required to recognize not only the correctness or incorrect-
ness of sequences that are parts of known ensembles, but also the
acceptability of new sequences. For such recognition, criteria must
be established for exclusion of the unacceptable. This effort will
examine current theories of structural linguistics in relation to
problems of automatic content processing of natural language text.
Criteria will be sought that can enable disqualifications on other
than *a priori* basis. The possibility will be explored of systematizing
graded criteria of unacceptability within the structure of correla-
tional grammar that will account for characteristics from various
levels of context. An attempt will be made to formalize procedures
for organizing semantic data into semantic factors so that the factors
can be accessible at all levels of correlation within a sentence. The
study will utilize factor data isolated by previously sponsored re-
search on correlational grammar.

Summary of Recent Research Results

Efficient semantic controls on correlations between words in a
sentence can have the practical result of improving the quality of
sentence analysis by (1) reducing the number of alternative overall
structures assigned to a sentence; (2) by reducing the number of
false correlations; (3) by clarifying relationships between words in a
sentence; and (4) by disambiguating individual words.

The task of establishing such controls is not easy. It seems clear
though that something less than an universal encyclopedia is needed
to do the job. The only practical way to discover semantically rele-
vant facts is to compare the interpretations yielded by the automatic
sentence analyzer with the interpretations accepted by a human

124

reader; by collecting the human readers explanations of his rejection of some interpretations; and by isolating the recurrent factors in these explanations. It seems worthwhile to go ahead far enough to find out how frequently the lack of specialized information actually in practice upsets the interpretation of a sentence. Experience so far suggests that it will not be frequent enough to undermine the usefulness of the procedure in practical applications.

Strong claims have been made for the ability of generative grammars to "assign" semantic descriptions to sentences, untouched by human intelligence. It would seem though that the *process* of generating a given sentence doesn't require the intervention of an intelligent person, but it would appear that the *specification* of the process does. This distinction has not been previously made. Yet, it is because of this difference that generative grammars do not directly lend themselves to adaptation for purposes of automatic sentence analysis; for they prescribe no way of using a sentence as input to an explicit system of rules which would design a description to that sentence.

In work in progress, a scheme is outlined for computing possible deep structures of a sentence in parallel with its surface structures by establishing systematically, for each correlational type, its characteristic implications for deep structure. Such a scheme should facilitate semantic controls and also yield very considerable simplifications in the existing reclassification system.

COMPUTABLE SEMANTIC DERIVATIONS AND THEIR EMPIRICAL BASIS IN LANGUAGE

Yorick Wilks

Institute for Formal Studies,
Los Angeles, California

Project Description

A number of different approaches are being studied in current research on natural language text toward the development of a mechanizable method of recognizing structure and meaning (semantics) for such purposes as machine translation and information storage and retrieval. This effort is for the extension of a study of fundamental semantic structures in continuous text by empirical methods. The system employs a finite but extendable inventory of coded semantic patterns called templates and represents text as an output concatenation of these patterns. The effort will explore resolution of word-sense ambiguity in paragraph pieces by means of rules of semantic sequence (both structural and content rules) and rules of compression (tree constructing rules) both of which operate on the templates. A non a priori procedure is being developed for constructing negation classes for rejecting nonapplicable patterns. The proposed analysis is expected to be formally analogous to a computational proof or derivation procedure in which the templates earlier in text order function as axioms and theorems and the resolution rules function as rules of proof. The analysis is also aimed at coping with limits of possible use of words in a language, by using the rules to *construct* the representation of a word-sense in certain cases.

Summary of Recent Research Results

The analysis is now being implemented in LISP 1.5 on SDC's Q-32 computer. Ten paragraphs are being analyzed, five of which are control texts from which structural information is derived to give content to the above mentioned rules. The other five are the experimental texts to which the rules will be applied to resolve their word sense ambiguity. As well as the rather strong application of the rules mentioned above, in which the order of application of the rules is specified in advance, a promising, weaker, approach is also being implemented in which the order is not so specified. This approach is in effect a semantic parser, that parses paragraphs semantically in a way analogous to the grammatical parsing of sentences.

126

Reports and Publications

Yorick Wilks, "Semantic Consistency in Text.. . . . an experiment and some suggestions" (System Development Corp., SP-2758, February, 1967).

Yorick Wilks, "Transformational Grammars Again" (System Development Corp., SP-2936, September, 1967).

127

5.10 AUTOMATIC PROGRAMMING

Bernard A. Galler

University of Michigan, Ann Arbor, Michigan

Project Description

Automatic programming concepts and techniques are being analyzed for systematic presentation. The analysis will include examination of the application of symbolic logic to language structure. Consequences suggested by the analysis will be explored, e.g., the automatic generation of translators for artifical languages; ability to change the definition of the language being translated, even during the translation; and possibliities of redefining operations and modes of arithmetic in procedure-oriented computer languages.

Summary of Recent Research Results

This grant supported research leading to the writing of a book by the principal investigator while on sabbatical leave from the University of Michigan. The research was done during the period from July 1, 1965 through April 14, 1967. During part of the period of the grant, the work was done in Amsterdam, Holland, and the remaining part of the work has been carried on at the University of Michigan, Ann Arbor, Michigan. Most of the research was done in collaboration with Professor Alan J. Perlis of the Carnegie Institute of Technology, who was also on sabbatical leave.

The major part of the research centered around a book which is tentatively titled "Algorithms, Definitions, and Data Structures." The principal ideas involved in the book are concerned with a programming approach to Markov algorithms, an alternative view of programming through ALGOL, a new proposal and development of the use and processing of definitions in a procedure-oriented language, a general treatment of data structures within the capability of a definition facility, and a treatment of binding time in a computing system. The topics just mentioned are roughly the contents of the individual chapters of the book. The first four chapters are quite finished, and have been put on multilith masters for use in a trial run which Professor Perlis has carried on through the 1966 academic year. The last chapter is under development.

128

The third item mentioned above, definitions in a procedure-oriented language, was the subject of a rather lengthy paper presented by the principal investigator and his colleague, Professor Perlis, at the 1966 National ACM Conference in Los Angeles. It appears as an article in the April 1967 issue of the *Communications of the ACM*.

Reports and Publications

Bernard A. Galler and Alan J. Perlis, *Comm. of the ACM*, "A Proposal for Definitions in ALGOL, **10**, Number 4, pp 204-219 (April 1967).

Bernard A. Galler, "Automatic Programming" (Final Report, April 1967).

ANALYSIS OF STRUCTURES OF
HUMAN COMMUNICATION SYSTEMS

Frederic T. Sommers, David Rothenberg,
and David M. Massie

New York Research Group Inc.,
New York, New York

Project Description

A similarity of structure has been found in models describing information transmission through aural perception and through the use of natural language. Prior research indicates that human systems organize classification and retrieval operations according to a structure that derives both from physiological limitations and the specific tasks that are performed. Continued research is directed toward an attempt to formulate a general theory of human communication systems that will exemplify the theories of aural perception and semantic structure as special cases. Experimental testing of a model for the perception of tones in context is in progress. The formal category (tree) theory was more simply axiomatized with the predication relation taken as primitive; it provides as formal analogy to natural language usage. Programs are being prepared for generating and analyzing language trees for large numbers of terms. Preliminary results verify an assumption that natural language uses sets of minimal ambiguity. Extensions of category theory aim toward mapping an entire vocabulary onto a formal semantic structure that shows the relative significance of all terms. A quantifier-free logic of monadic predicates was extended to binary (relational) predicates; the aim is extension to n-ary relations.

Summary of Recent Research Results

The model for the perception of tones has been extended to apply to a wider range of cases. Computations have been made in application of the model to particular cases, and combinatorial techniques developed for computer use. Extensive ethnomusicological verification of the model has been obtained. Programs for computer generation of semantic structures have been prepared and are being used to investigate tree structures and ambiguity resolution for natural language. The quantifier-free logic of predicates has been extended to n-ary relations and to compound predicates, and an algebraic model for the theory is being developed. A widely

130

applicable formal learning theory has been formulated which is a generalization of the theories of aural perception and of semantic structures.

The theory provides a formal model of learning process and is designed for computer application to specific learning tasks. Learning is treated as a process of data structuring, the structures determined by the task to be performed. The computer realization of the model will print out a representation of the resulting information structures, which are models of the learning tasks performed. The theory promises to have great generality and wide application.

Reports and Publications

David Massie, "Logic of Contraries and Denial" (Technical Report No. 5, to be published).

David Rothenberg, 'A Mathematical Model for the Perception of Tones in Context" (Technical Report No. 1, Dec. 1966; AFOSR 67-0415, AD647270).

David Rothenberg, "A Mathematical Model for the Perception of Tones in Context, Part II" (Technical Report No. 2, May 1967; AFOSR 67-1489, AD655117);

Frederic T. Sommers and David Massie, *Bulletin of the American Mathematical Society*, "Provability and Predicability" (September, 1966).

Frederic T. Sommers, "On a Fregean Dogma" in *Problems in the Philosophy of Mathematics* (The North-Holland Publishing Co., 1967).

Frederic T. Sommers and David Massie, "A Theory of Semantic Structure for Natural Language" (Technical Report No. 3, to be published).

Frederic T. Sommers, "Quantifier-Free Logic of Predicates" (Technical Report No. 4, to be published).

5.12 AUTOMATIC ENGLISH SENTENCE ANALYSIS

Ernst von Glasersfeld

PEC, Inc., Athens, Georgia

Contract Number F44620-67-C-0028

This research project is described in 5.7. The present grant is with
the Georgia Institute for Research, Inc., Athens, Georgia.

132

5.13   PROBLEMS IN MACHINE PROBLEM SOLVING

Saul Amarel

Radio Corp of America, Princeton, New Jersey

Project Description

A basic understanding of how to analyze and represent various types of problems is an essential first step to developing efficient and effective procedures for problem solving by machine. Prior research on this effort examined theory formation processes, heuristic problem solving procedures, and methods for realizing them. Results have yielded a conceptual framework for ordering a variety of problem types for machine solution. Continued research will exploit the distinction between derivation and formation problems, i.e., derivation proceeding mainly from functional conditions to be satisfied to the structure of the solution, and formation proceeding mainly from structural to functional properties of solutions. A promising approach to formation problems has been to find a mode of representing candidate solution structures that establishes a meaningful relationship between them and functional aspects of the solution, thus enabling reformulation of the formation problem to resemble a derivation problem amenable to known solution procedures. Attention will focus on the   problem of representation and such associated questions as its relationship to problem solving power and similarities between transitions of representations (i.e., automatic change in a machine point of view) and creative processes. Continued study of question-answering problems will emphasize relationships between the languages involved and the efficiency of answer generating procedures. Research on linguistic aspects of problem solving will concentrate on formalizations of language translation processes.

Summary of Recent Research Results

Considerable progress has been made towards clarifying the problem of representation for derivation problems. A sequence of three reduction procedures has been formulated, each more powerful than the previous one, for constructing a simplest proof to a theorem in the propositional calculus. The best procedure in the sequence is based on a novel graph representation of proofs which induces an efficient global approach to proof construction. Study

133

has shown that two main types of representation changes are at play in transitions between these procedures; one has the nature of a homomorphism, and the other of language translation. Similar results were obtained in detailed studies of problems of reasoning about actions. The homomorphic transitions are based on the exploitation of acquired problem-specific knowledge, such as redundancies, symmetries, and critical points in the search space. Transitions of the language translation type were found to be especially sensitive to choices of primitive objects and predicates for descriptions of situations.

A previously developed natural inference framework for derivation problems was successfully applied to the construction of models for syntactic analysis. These models shed light on key features of many existing parsers, and have helped in the synthesis of more powerful ones. The natural inference framework has proved to an excellent basis for the study of question answering systems. Alternative descriptions for graphical data bases have also been developed and analyzed. Results on completeness and deductive ease of descriptions have contributed further to an understanding of representations in problem solving.

The main results in program formation have been in the extension of the descriptive mechanism for programs. Study has shown that the relationship between the language of programs and its corresponding (previously obtained) algebraic model can be represented as a language translation.

Research interest in design problems has also increased primarily due to their frequency in real-life situations (Design problems are mixtures of formation and derivation problems). By considering the architectural problem of designing floor plans, considerable progress has been made in the static-linguistic aspects of describing solutions in the process of evolution, and also in the dynamic-strategic aspects of controlling such an evolution.

Reports and Publications

Saul Amarel, "An Approach to Heuristic Problem Solving and Theorem Proving in the Propositional Calculus", to appear in *Proceedings of the Conference on Computer Science and Systems,* University of Western Ontario, September 1965 (University of Toronto Press, 1967).

Saul Amarel, "Problem Solving Procedures for Efficient Syntactic Analysis", in *Proceedings of the ACM 20th National Conference,* Cleveland, Ohio, August 1965.

134

Saul Amarel, "More on Representations of the Monkey Problem" (Carnegie Institute of Technology Seminar Notes, CS-691, March 1966).

Saul Amarel, *IEEE Spectrum*, "On the Mechanization of Creative Processes", April 1966.

Saul Amarel, "On Machine Representation of Problems of Reasoning About Actions–The Missionaries and Cannibals Problem" (RCA Laboratories and Carnegie Institute of Technology Report, June 1966).

Saul Amarel, Opening Remarks in Session on "Representations and Modeling in Problem Solving" and talk on "Future Directions in Problem Solving Systems" to appear in *Proceedings of the AFOSR Conference on Intelligence and Intelligent Systems*, University of Georgia, Athens, Georgia, 1967.

Saul Amarel, "On Representations of Problems of Reasoning About Actions", to appear in *Machine Intelligence 3* (University of Edinburgh Press, 1968).

M. Kochen, "On the Representation of Limited Information by Means of Pictures, Trees, and English-like Sentences" (RCA Laboratories and University of Michigan Mental Health Institute Report, September, 1965).

M. Kochen, "Translation of English-like Queries into Efficient Computer Search Programs for Question Answering" (RCA Laboratories and University of Michigan Mental Health Institute Report, November 1965).

M. Kochen, "Translation of English-like Questions about Simple Pictures in Flow Diagrams to Produce Answers by Computer", submitted to the Journal of the ACM.

F. Larkin, "Complexity in Program Translation", to appear in the *Proceedings of the Symposium on Logic, Computability and Automata Theory*, Rome, N.Y., August 1965 (Thompson Books, 1967).

D.A. Walters, "An Analysis of an Approximate Sequential Procedure for Interval Estimation" (Scientific Report Number 1, July 1963; AFOSR 5155, AD438559).

R.O. Winder, "An Evaluation of Heuristics for Threshold Function Test-Synthesis (RCA Laboratories Report, December 1965).

Invited Lectures

S. Amarel, "Automatic Formation of Computer Programs by Methods of Successive Approximations"; presented at the Monthly Seminar on Automatic Computers and their Capabilities, Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia (March 4, 1964).

S. Amarel, "Research on Automatic Theory Formation"; presented at Bell Telephone Laboratories Computer Seminar, Murray Hill, N.J., (April 14, 1964); Colloquium in Computer Science, Columbia University Computer Center, N.Y., (May 1, 1964); and Graduate Seminar in Artificial Intelligence, Psychology Department, Princeton University, (May 19, 1964).

S. Amarel, "Problem Solving and Theory Formation by Machine"; 8 lectures presented at the Special Summer Session on Advanced Theory and Practice of Digital Computer Programming, Moore School of Electrical Engineering, University of Pennsylvania, (June 29-July 10, 1964).

S. Amarel, "On the Formation of the Concept of a Transformation by Computer"; presented at the 1964 International Congress for Logic Methodology and Philosophy of Science, Jerusalem, Israel, (Aug. 26, 1964).

S. Amarel, "Intelligent Problem Solving by Computer"; presented at RCA Laboratories Colloquium, (Dec. 21, 1964).

S. Amarel, "Languages of Representation in Problem-Solving Procedures of the Formation Type"; presented at the Symposium on Logic, Computability and Automata Theory, held, in Rome, N.Y. (August 25 to 27, 1965); summary to appear in Proceedings of the Symposium, Thompson Books, (1967).

S. Amarel, "Comments on the Mechanization of Creative Processes"; contributed in a panel discussion at the IFIP Congress '65 in New York (May 28, 1965).

S. Amarel, "Artificial Intelligence"; presented at seminar in Advanced Computer Techniques, Carnegie Institute of Technology, Pittsburgh, Pa., (April 5, 1965).

S. Amarel, "Problems of Representation in Problem Solving Procedures"; presented at the Systems and Communications Sciences Colloquium, Carnegie Institute of Technology, Pittsburgh, Pa., (April 6, 1965).

S. Amarel, "On Two Types of Intelligent Problem Solving Procedures"; presented at Systems Research Center Seminar, Case Institute of Technology, Cleveland, Ohio, (April 7, 1967); and Electrical Engineering Seminar, Purdue University, Lafayette, Indiana, the talk was telecast and taped, (April 8, 1965).

S. Amarel, "Invited Critique of paper by H. Borko on "The Conceptual Foundations of Information Systems"; presented at 8th Annual Summer Symposium of the Syracuse University School of Library Science, Theme: The Foundations of Access to Knowledge, (July 28, 1965).

S. Amarel, "Problem Solving by Computer"; presented at The 2nd of the Invited Lecture Series of 1965-66 at the Systems Research Center, Carnegie Institute of Technology, (Dec. 3, 1965); and Seminar of Electrical Engineering Department, University of California, Berkeley, (Dec. 8, 1965).

136

S. Amarel, "Problems in the Design of Problem Solving Procedures";
presented at Computer Science Colloquium, Stanford University,
(Dec. 6, 1965).

S. Amarel, "Automatic Theory Formation"; presented at Computer
Science Department, Stanford University, (Dec. 7, 1965).

S. Amarel, "Complex Information Processing Systems"; graduate
course (CS-654-A) given at the Computer Science Department,
Carnegie Institute of Technology, Spring Semester, (1966).

S. Amarel and Alan Newell, "Advanced Seminar in Computer
Science—On Problem Representations"; (CS-691); given at the
Computer Science Department, Carnegie Institute of Technology,
Spring Semester, (1966).

S. Amarel, "Representations and Modeling in a Program Formation
Problem"; presented at the Research Seminar of the Mental
Health Research Institute, University of Michigan, (Dec. 2, 1966).

S. Amarel, "Research Problems in Problem Solving by Computer";
presented at the Seminar on Computers and Automata, Poly-
technic Institute of Brooklyn (Feb. 17, 1967).

S. Amarel, "The Current Status of Artificial Intelligence"; presented
at the monthly meeting of the Columbia University Seminar on
Relations Between Research and Education and Computers,
March 20, 1967.

S. Amarel, "Problem Solving by Computer - State of the Art and
Open Problems"; 3rd lecture of a Columbia University Colloquium
Series at the School of Engineering and Applied Science, March
24, 1967.

S. Amarel, "On the Problem of Representation in Computer Problem
Solving"; presented at the Digital System Seminar, Princeton
University, (April 13, 1967).

5.14 SURVEY OF COMPUTER LANGUAGES FOR SYMBOLIC AND ALGEBRAIC MANIPULATION

Bertram Raphael

Stanford Research Institute,
Menlo Park, California

Project Description

Computers are being used increasingly to do such non-numerical tasks as theorem proving, work in pure mathematics (group theory, topology, analytic number theory), formal algebraic manipulation, and the formulation of techniques for solving problems (list processing, string handling). The volume of activity suggests the desirability of providing a perspective through a survey of the computer languages that have been developed for these symbolic and algebraic manipulation operations. This effort is for partial sponsorship of an initial survey of about fifteen languages in current use. Principal features of each language will be described, and the languages will be compared with respect to such factors as application areas, convenience of use, environmental constraints, and various aspects of over-all performance. The survey should facilitate knowledgeable selection of a computer language for future studies exploring computer applications in non-numerical mathematics.

Summary of Recent Research Results

This research effort resulted in a survey report on 18 computer languages used for symbolic and algebraic manipulation. These languages are: ALTRAN, AMBIT, COGENT, COMIT, CONVERT, CORAL, DYSTAL, FLIP, FORMAC, FORMULA ALGOL, IPL-V, LISP 1.5, LISP 2, L^6, PANON, SLIP, SNOBOL, and TRAC. Several other languages are also mentioned briefly in the report. For each of six groups into wich the languages are classified, the report (a) describes properties that members of the group have in common, (b) gives a brief description of each language in the group, including an excerpt from a program in the language that demonstrates the kind of problem for which the language is well suited; and (c) briefly compares the features of the languages in the group. The report contains three appendices: (1) a reference chart that summarizes the features of all of the languages; (2) a comparison chart that emphasizes the salient distinctions between selected pairs of similar languages; and (3) a set of annotated examples of programs in various languages that

138

solve similar problems, thus illustrating the differences in data representations, program forms, and notations.

Reports and Publications

Bertram Raphael, "Survey of Computer Languages for Symbolic and Algebraic Manipulations" (Final Technical Report, March 1967; AFOSR 67-0811, AD649401).

5.15   A THEORY OF COMPUTER LANGUAGES

Andrzej Ehrenfeucht

Stanford University, Stanford, California

Project Description

Computer languages must be read both by men and machines. Much current research on artificial languages is directed toward the construction of precise languages which can be readily learned and read by man. This effort studied the structure and usefulness of languages between the extremes of Algol-like generality and machine-language specifity. An attempt was made to develop a theory of languages for describing computational processes and the organization of computations that can be employed in the design and implementation of man-manipulable machine languages. Previous results concerning the interpretation of bracket-free formulas used in composing functions will be formalized and extended to languages used for recursively defined processes. Applications of this theory will be sought to the formulation of useful procedures in such areas as mechanical theorem proving.

Summary of Research Results

Two reports have been prepared under this grant. The first is entitled "Practical Decidability" and deals with the complexity of problems measured by the size of memory of the computer that is necessary for solving them. The main results are: (1) There are problems for which complexity is an exponential function of the length of the data (measured by number of digits used for coding); (2) *The complexity of the above type cannot be improved* by any change in the organization of the computer, change of the programming language, etc. . . ; (3) such problems can easily be formulated within some formal mathematical system that is Peano Arithmetic.

The second report is entitled, "A Theorem Concerning Linear Languages." The paper deals with the existence of a finite kernel of a language; that is, the finite set of minimal sentences (it is known that natural language has finite kernels). The main result is that under some general conditions artificial languages also have finite kernels.

Reports and Publications

Andrzej Ehrenfeucht, "Practical Decidability" (Internal Stanford University Report, Philosophy Department).

Andrzej Ehrenfeucht, "A Theorem Concerning Linear Languages" (Internal Stanford University Report, Philosophy Department).

141

5.16   **THEORY OF PROGRAMMING
ALGORITHMIC LANGUAGES**

Seymour Ginsburg

System Development Corporation,
Santa Monica, California

Project Description

Data processing by machine presupposes the existence of a
programming language with which to describe the data flow for
problem solution. Programming language construction has been
pragmatic. The design of many user-oriented languages is directed
toward natural-language constructs that are sufficiently rich to
permit the description of a range of problems and sufficiently
formal to permit efficient translation to machine languages. This
effort seeks to develop a theory of programming languages that
can enhance existing languages and aid the design of new ones.
Research includes the formulation of more accurate models of
programming languages, examining theoretical implications of
the models, particularly for computer processing, studying the
effects of various computers on the languages, determining the
existence of decision procedures for problems of interest about
programming languages, and identifying languages for which
compilation can be readily effected. Basic tools of discrete
mathematics and abstract algebra are being employed toward
establishing such factors as the limitations on higher-order
languages, laws that apply to language structure, and techniques
for synthesizing and analyzing the languages.

Summary of Recent Research Results

Several facets of context-free languages (CFL) and arbitrary
rewriting systems were investigated. A number of problems about
CFL were shown to be unsolvable even in the presence of partial
information (1). A new device was introduced which recognized
exactly the CFL (2). Unambiguity of general rewriting systems was
defined and discussed (4).

A device was introduced (3) which had its memory organized as
a list. Special attention was then given to the recognizer and
associated language which resulted when the input was read one-
way and in real-time.

142

In order to unify the treatment of many recognizers and algorithmic languages in the literature, the notion of an abstract family of languages (AFL) was defined (5). Many of the standard properties for diverse formal languages were shown to hold for AFL. In addition, intimate connections with recognizers and transducers were established. Recent results lend weight to the belief that AFL are to play a major role in the theory of formal languages.

Reports and Publications

J. Ullian, "Partial Algorithm Problems for Context Free Languages", 10 October 1966, 34pp., TM-738/027/00.

T.N. Hibbard, "A Generalization of Context Free Determinism" 21 November 1966, 65 pp., TM-738/028/00.

S. Ginsburg, and M.A. Harrison, "One-Way Real-Time List-Storage Languages. 3 January 1967, 44 pp., TM-738/029/00.

T.N. Hibbard, "The Equivalence of Context Limited Grammars to Context Free Grammars." 16 January 1967, 63 pp., TM-738/030/00.

S.Ginsburg, and S. Greibach, "Abstract Families of Languages", 17 April 1967, 57 pp., TM-738/031/00.

S. Ginsburg, and G.F. Rose, "A Note on the Preservation of Languages by Transducers, 23 May 1967, 8 pp., TM-738/032/00.

5.17 STATUS OF COMPUTATIONAL LINGUISTICS AND RELATION TO MACHINE TRANSLATION

Winifred P. Lehmann

University of Texas, Austin, Texas

Project Description

Controversy, highlighted recently by NRC Publication 1416, *Language and Machines*, has tended to suggest excessive federal expenditure on machine translation research. Recommendations emphasize the need for increased research in computational and descriptive linguistics as well as machine aids for translation. Considerable linguistics research that has been attributed to machine translation has been concerned with computational and descriptive linguistics, but no current state-of-the-art report exists for this literature. This effort will survey the present status of computational linguistics and relate it to machine translation research. The study will consider the languages that have been examined, the merits and shortcomings of various methods, software and hardware aspects of the automatic language analysis problem, and achievements to date. Results are expected to clarify the record with respect to research emphasis directly oriented to machine translation and research relevant to it as well as provide a current assessment of existing knowledge as a basis for future research. A by-product will be a bibliography in machine-retrievable form to an extensive reference collection in computational linguistics.

Summary of Recent Research Results

Results of this study will be published in book form, under the title: *Machine Translation Today*. After an introduction, stating the presuppositions leading to the assumption that machine translation might be feasible, the work that has been undertaken is discussed in three sections. In these sections the MT systems that have been developed and put into use are described: 1) MT systems based primarily on syntactic descriptions. 2) MT systems based on sematic as well as syntactic descriptions. 3) Long-range research programs. The book is compact, restricted to MT research which may be typical of possible approaches. Concentrating on the theories underlying such research, it does not provide a full statement on every research project

144

that has been undertaken. The underlying aim is to present a statement which will provide for non-specialists a general view of the undertakings, accomplishments, and future prospects of work in MT. A final chapter discusses prospects. In a dynamic field like MT and Computational Linguistics, a constantly up-dated bibliography is essential. The project has permitted the establishment of a comprehensive bibliography in machine-retrievable form. In keeping this bibliography up-to-date, editorial principles must be applied to remove obsolete materials and add only appropriate new materials.

COMPARATIVE ANALYSIS OF FOUR LANGUAGES IN RELATION TO YNGVE's DEPTH HYPOTHESIS

Ernesto Zierer

Universidad Nacional de
Trujillo, Trujillo, Peru

Project Description

The design of synthetic languages approximating natural languages and the machine manipulation of natural languages for such purposes as abstracting and translation depend on algorithms that can account for sentence structure and contextual meaning. This effort proposes to explore the Yngve depth hypothesis concerning certain limitations on the structural properties of language, i.e., on the order in which constituents may be expanded. Basic sentence patterns in four languages will be examined: English, Spanish, German, and Japanese. The patterns will be related to the communication patterns they disclose, i.e., the information content and relationships the sentences express. The sentence patterns for the four languages will be compared and analyzed statistically with respect to their maximum depth and their depth load, the latter being a criterion for measuring the quantitative concentration of depth in a syntactic construction. Features characterizing each language will be determined which should indicate approaches to and limitations on their machine manipulation.

Summary of Recent Research Results

The final report for this project gives the results obtained in comparing English, German, Spanish and Japanese in the light of Yngve's Depth Hypothesis. Sixty-eight different language patterns distributed over 111 different sentences in these languages and a scientific text are analyzed for maximum depth and depth load. The patterns are not *sentence* patterns (i.e. structural patterns typical of one of the 4 languages) but rather *communication patterns*. The basic idea of the Depth Hypothesis is that producing an utterance that aims at a communicative effect implies the incurring of commitments that are imposed by the grammar and which may or may not be due to psycholinguistic restrictions that limits the number of commitments at any one time. Findings in this project includes the following: a) There is a limit to regressive expansion language analyzed; b) Japanese stands

146

· stinctively apart from the three Indo-European languages with a
considerably greater depth; c) Specific syntactic devices responsible
for greater depth in Japanese include end position of the verb—
adjective + noun + post-position contructions — pronoun + particle
constructions — nesting — interrogative particle at the end of the
sentence — noun attribute + particle + modified element construction
— sentence attribute functioning as a relative clause before the
modified element — and subordinator after the subordinate clause.
d) The concept for depth load is a criterion for further differentiation
of the syntax of the 4 languages. e) Regressive structures often
reflect a *collectivistic* mentality vs. and *individualistic* one as ex-
pressed in progressive structures. f) The syntactic devices producing
structures of greater depth are a major problem for the learner.
Mastery by special drills graded according to increasing depth is
suggested.

Reports and Publications

Ernesto Zierer, Kisang Lee and Anibal Sanchez, "A Comparative
   Study of Yngve's Depth Hypothesis" (Final Technical Report,
   May 1967; AFOSR 67-1542, AD659441

147

LANGUAGES FOR AUTOMATIC
PROGRAMMING TOOLS

Alfonso Caracciolo di Forino

Universita degli studi di Pisa, Facolta di Scienze Matematiche,
Pisa, Italy

Project Description

A programming language must be able to express both how to
define a job (i.e., the set of operations a man or machine is to do)
and how to execute it (e.g., by a set of machine commands). Although
much practical knowledge has been accumulated in designing
programming languages and systems for particular problem-solving
and process control tasks, many aspects of programming language
theory that could aid in programming language design, particularly
for complex dynamic systems, have not been formulated or explored.
This effort seeks to examine, cojointly, theoretical and practical
aspects of programming language theory and specification. Studies
are aimed at extending automata theory to describe realistic opera-
tions of practical machine tools. Additionally, starting from a for-
mal, rigorous definition of languages for defining geometric entities,
requirements will be sought for problem-oriented languages that can
describe machine operations in terms intelligible to a job-executor
(man or machine). The research is expected to add to an understand-
ing of fundamentals of the pragmatics of language, i.e., those re-
lationships that exist between signs and symbols and those who pro-
duce and use them.

Summary of Recent Research Results

Research has progressed in three directions as outlined in this
project's general description.

1. Extension of the automata theory approach outlined in (1)
   to a more realistic description of a Machine Tool. A model has
   been developed that takes into account the fact that usually
   only a portion of the cutting tool has an effective cutting sur-
   face. Therefore, a condition must be added forbidding any
   intersection of the non-cutting part of the tool with the work-
   piece. Moreover, a model is being investigated that considers
   speed of motion of the tool which plays an essential role in the
   actual behavior of a machine tool.

2. The vigorous definition of languages for defining geometric forms requires an investigation of the basic syntactic features of declarative languages, i.e., languages containing definitional statements that should be constructed as local metasyntactic rules. An attempt is being made to find adequate grammatical models for these languages and for solving the general problem described in (2).

3. An effort is being made to find an algorithm to determine an optimal path for a cutting tool that removes the maximum allowable quantity of material at each step. We plan to incorporate both technological and geometric parameters into the algorithm.

Reports and Publications

Alfonso Caracciolo di Forino, Calcolo, "Linguaggi Programmativi Speciali", 2, Supplemento 2, pp. 69-89 (1965; English translation, Lito-offset Felici, Pisa, Italy).
Alfonso Caracciolo di Forino, "On the Concept of Formal Linguistic Systems", in *Proceedings of the IFIP's Working Conference on Formal Language Description*, T.B. Steel, Ed. (North-Holland, 1966) pp. 37-51.

5.20 **GENERALIZATION AND EXTENSION OF PROGRAMMING LANGUAGE CONCEPTS**

Philip R. Bagley

University City Science Center, Philadelphia, Pennsylvania

Project Description

A major problem in the field of computer programming is the large amount of time and effort necessary to proceed from a point where the programmer has a mental concept of that which he wishes his program to accomplish to the point where his program has been written, debugged, and is operational. This effort investigates the possibility of separating the task of writing a program into two distinct phases. The first phase would be to write the program in an algorithmic or logical language, independent of the characteristics of the specific computer being used. The second phase, which can be replaced by a mechanical process, would be to specify in an implementation language the manner in which the program will be executed on a specific computer. Initially, methods will be sought to construct an algorithmic language by defining and representing a broadly applicable set of basic data elements and structures, and a set of basic processes for their manipulation. A great reduction in time and effort may be realized when modifying an operational program or when converting from one computer system to another.

Summary of Recent Research Results

Research to date has been primarily concerned with the question: "What are the concepts, independent of specific machines, that systems analysts use in framing procedures to be carried out on machines?"

Research phase one was a survey of existing programming languages. Phase two was the development of a frame work for language description into which could fit the parts as they evolve during the research. The Major parts of this framework are data descriptions, data transformations, and sequencing rules. Phase three was the development of requirements for data description, both of atomic data elements, such as variables, and of composite data elements, such as a complex numbers. Phase four outlined how data elements could be represented in a hypothetical machine.

150

By-products of this effort have been: 1) A bibliography of over 300 items related to the design and development of programming languages. 2) A wide variety of simple problems, to be used as a "proving ground" for the language being developed. 3) A disproof of the original conjecture that a directed graph is an appropriate generalization of all other data structures.

Section 6

THEORETICAL FOUNDATIONS OF
INFORMATION SCIENCES

Most of the studies in this task concern developments in symbolic logic. We were not thinking of the new mathematics called for by von Neumann to describe intelligent activities when we established the task. Our sponsored projects in multivalued logics increasingly suggest this as a possible route. Since the McCulloch-Pitts logical calculus of 1943, procedures have been sought for expressing such "ideas immanent in nervous activity" as learning and recognition and the ability of automata to function reliably under duress, with unreliable components, and with incomplete information. Space flight and hostile environments now impose requirements on logical elements of adaption and self-repair.

Automata theory is another route for building intelligent systems. The use of algebraic methods to describe the language and behavior of automata may lead to machines that are both fundamental and behavioristically simple. Combinatorial algebras may give useful models for information retrieval systems. Techniques for simplifying proof procedures offer possibilities of eliminating exhaustive enumerations that are time consuming and can exceed machine capacities. Questions concerning algorithmic unsolvability and other problems in recursive function theory (the theory of computibility) are amplifying notions of constructibility, decidability, consistency, and completeness.

The logic models supply formal approaches that can be applied to problems in the structuring and processing of information and language. The precise descriptions of automata are models of realizable networks and define bounds on their performance. Automata are beginning to be assessed for the solution of natural language problems. However, it is expected that applications to information problems will await further examination of such questions as complexity, equivalence, and the behavior of automata with and without restrictions on time and space.

6.1 ALGEBRAIC METHODS IN AUTOMATA THEORY

Marcel P. Schutzenberger and Maurice P. Nivat, Association
Pour L'Etude Et La Recherche En Traitement
De L'Information, Paris, France

Current Contract Number AF-61 (052) 00945

Project Description

The use of algebraic methods in automata theory is leading to
descriptions of automata (machines) which are both fundamental
and behavioristically simple. Recent results include a theorem on
group free automata applicable to loop-free nets via a simple class
of predicative formulas; a theory of incompletely specified automata
which provides solutions to the general problem of approximating
any event by some regular one; and a theory of variable-length codes.
This effort will continue research toward the development of a
theory of finite transducers; a theory of abelian languages; and a
theory of codes. A definition of transducer (an automaton which
provides a standard form for a word of a given language) is enabling
the description of T-languages which form an intermediate class
between unambiguous and deterministic languages. Abelian languages
provide a specific link between the theory of languages based on an
abstract alphabet and the so-called Pressburger predicates which,
being strictly arithmetic, seem adequate for the theory of numerical
automata. Several basic conjectures in the theory of codes will be
examined for a better understanding of combinatorial problems and
the meaning of so-called unique decipherability problems.

Summary of Recent Research Results

T ree closely interrelated topics have been developed in the course
of this research project. These are:

1. The development of various algebraic properties of monoids.
Results include:

   a. The proof of conditions for two words in a free monoid to
be powers of the same word (generalizing the results of Lyndon) and

   b. The proof of conditions on a set of positive integers to be the
lengths of the words in a synchronizing prefix code (thus, answering
a question raised by Wolfowitz on monoids of non-negative matrices).

154

2. The discovery of several properties of rational and algebraic subsets of monoids. For example, a complete characterization is given of the rational subsets of any commutative monoid. Also, progress has been made toward answering questions concerning the commutative closure of context-free languages. The use of transductions (defined as the rational subsets of the cartesian product of two monoids) leads to powerful theorems that cover classical results on context-free languages. The same tool provides a very general scheme for syntactic analysis of context-free languages and aids in proving many equivalence relationships of push-down automata (such as those of Ginsburg, Griebach, and Schutzenberger).

3. An analysis of the computing power of finite automata or nets of finite automata. A large number of computations on graphs can be carried out by simultaneously using a finite automaton to process the incidence matrix of the graph and several isomorphic finite automata attached to each vertex of the graph.

Reports and Publications

J. Berstel, *C.R.A.S.*, "Resolution pas un Reseau d' Automates du Probleme des Arborescences dans un Graphe", **264**, pp.388-390 (Paris, 1967).

M. Nivat, "Transductions des Languages de Chomsky", in preparation, to be submitted for publication to *Annales de l' Institut Fourier.*

M.Nivat, "Conguences de Thue et t-Languages, submitted to *Publ. Math. Inst. Hungarian Acad. Sci.*

M.P. Schutzenberger, *Journal of Combinatorial Theory*, "On a Question Concerning Certain Free Submonoids", **1**, pp. 437-442 (1966).

M.P. Schutzenberger and M.P. Nivat, *C.R.A.S.*, "Sur les Produits Semi-Directs Droits de Monoides", **263**, pp. 659-660 (Paris 1966).

M.P. Schutzenberger, "On Computable Sets of Numbers, in preparation, to be submitted for publication to *J. ACM.*

M.P. Schutzenberger, "On Synchronizing Properties of Prefix Codes", in preparation, to be submitted for publication to *Information and Control.*

M.P. Schutzenberger and A. Lentin, "A Combinatorial Problem in the Theory of Free Monoids", at the *Symposium in Combinatorial Theory*, Chapel Hill, 1967.

M.P. Schutzenberger and S. Eilenberg, "On Rational Sets in Commutative Monoids", submitted for publication to *Trans. Amer. Math. Soc.*

M.P. Schutzenberger, "Sur Certains Produits de Matrices non Negatives", submitted for publication to *Mathematische Zeitschrift.*

155

6.2 COMBINATORIAL ALGEBRA WITH APPLICATIONS TO INFORMATION SCIENCE

Frank B. Cannonito

University of California, Irvine, California

Project Description

Current information retrieval systems are primarily extrapolations of human intuition about the relationship of the use of information tags by which it can be retrieved. Recent analytic studies suggest that the application of mathematical theories to large classes of information retrieval problems may yield uniform methods of treating them and representing them with maximum economy for machine processing. This effort involves an application of group-theoretic concepts to information retrieval. The descriptors associated with a stored document can be regarded as defining a word on the total alphabet of descriptors. Since the order of the descriptors. Since the order of the descriptors is unimportant and no descriptor appears more than once, it is possible to regard the totality of such words as defining elements in a commutative monoid with idempotent generators. Solving the word problem in the associated monoid is viewed as the mathematical counterpart of the following information retrieval problem: given a set of descriptors for a collection of documents, how can certain combinations of descriptors be algorithmically replaced by equivalent combinations that describe related, relevant properties. Several aspects of this question will be examined.

Summary of Recent Research Results

The effect on solving the word problem in a group caused by giving the group its Cayley-representation as a group of permutations has been studied. This problem seems to be related to an apparently unsolved problem: is the group generated by the primitive recursive permutations the full group R of recursive permutations? We have succeeded in showing that R is not finitely generated by exhibiting a nonstationary infinite tower of subgroups of R, viz. R_3 R_4 ... R where for $\alpha \geq 3$, R_α is the group of permutations f such that both f and its inverse f^{-1} are at level α of the Grzegorczyk hierarchy. In addition we have obtained a sharper version of Julia Robinson's representation of (singularly recursive functions; thus, we can show each singularly recursive function f is representable as $f = AB^{-1} C$ where A and C are *fixed* elementary functions and B is an elementary permutation. These results are announced in the October 1967 notices of the AMS and will be submitted to the Proceedings of the AMS.

156

6.3 CONFERENCE ON INTELLIGENCE AND INTELLIGENT SYSTEMS

Robert A. McRorie and Thomas W. Mahler

University of Georgia, Athens, Georgia

Project Description

The Macy Conferences on (Circular, Causal, and Feedback Mechanisms) in the 1950's initiated and set a pattern for group discussions on topics concerned with intelligence and intelligent systems. Subsequent meetings have enabled consideration of specific topics in this area, but the field has developed sufficiently to warrant a conference in the nature of a Macy Conference to assess present research and consider future directions. This effort undertakes such a conference. It will be invitational and will differ from the Macy Conferences in that it is intended to permit established experts to interact with younger researchers who have begun to make fundamental contributions to the field. Unclassified proceedings with no distribution limitations are planned.

Summary of Recent Research Results

This Conference was held at the University of Georgia on January 16-18, 1967. It provided a forum for the review of many of the approaches that are being developed toward the design of intelligent machines. There were six sessions addressed to the topic of: foundational considerations; representations and modeling in problem solving; approaches to machine intelligence; and the structure of human and machine intelligence. Discussions ranged from the consideration of implications from human problem-solving protocols to the performance of robots. Requirements for language systems that reflect characteristics of intelligent behavior were also examined. The Conference underscored that the state of the art of artificial intelligence is in its nascency with respect to models and machine systems. Contributions from psychologists and psychiatrists suggested sources of experiential data that may prove useful in future research to mathematicians, linguists, and engineers.

157

6.4   RELATIONS BETWEEN LOGICAL STRUCTURE, TIME, AND NATURAL NUMBERS

Gotthard Gunther

University of Illinois, Urbana, Illinois

Project Description

Research on adaptive, self-organizing, and self-referential systems
that has been facilitated in recent years by the data-processing capa-
bilities of computers has been hampered by the lack of adequate
mathematics for describing observed or hypothesized phenomena.
This effort postulates the need for a trans-classic logic and a bridge
between this logic and mathematics to enable the formalization of
models of thinking self-referential systems. Prior research has yielded
clarification of concepts and a pattern for a kenogrammatic structure
that is expected to serve as a basis for a formalization of a theory of
ontological loci that can then furnish the missing link between
ontology and computer logic. Exploration of a system is proposed in
which all trans-classic values of logic are located outside the range
of 0 and 1 and form, together with two-valued classic logic, a place
value system of an ascending order of ontologies and logic structures
of increasing functional complexity. The only type of order that is
admitted for the theory of logical values is of the character of aleph
sub-zero, i.e., the order of the natural numbers. Relations between
quantity and structure will be examined.

Summary of Recent Research Results

During the second half of 1966, the initial stage of the general
research project which concerns new theoretical foundations of
Science - and especially cybernetics - was completed. Since then it
was discovered that the kenogrammatic structures which is sub-
divided into proto-, deutero- and trito-structure permits a secondary
subdivision of trito-structure. Up to this time it was considered un-
resolvable. However, it is possible to subdivide into trito-structure
proper and morpho-grammatic structure. This subdivision had a
peculiar effect: If trito-structure and morpho-grammatic structure
are considered a unit, the whole kenogrammatic structure is basically
a logical context. If it is resolved into two parts, the kenogrammatic
structure partly loses logical significance and develops into a *purely*

158

mathematical context of relations. From this discovery originated the idea of a trans-classical system of mathematics. In pursuit of this idea an analysis of the theory of natural numbers was undertaken. It was found that natural numbers could be built up on the basis of proto-, deutero- and trito-structure. It followed that an increase in quantity is interpretable as a parallel increase in logical complexity. This is contrary to present trends of thought where a mere quantitative increase is not supposed to imply an increase in logical structure.

Reports and Publications

Gotthard Gunther, *Zeitschr. f. Philos. Forschung,* "Die Aristotelische Logik des Seins und die Nicht-Aristotelische Logik der Reflexion" (The Aristotelian Logic of Being and the Non-Aristotelian Logic of Reflexion), **12,** pp. 360-407 (1958).

Gotthard Gunther, *Augenblick,* "Die Gebrochene Relationalitat" (The Refracted Rationality), **3,** pp. 1-26 (1958).

Gotthard Gunther, "Idee und Grundriss einer Nicht-Aristotelischen Logik" (Idea and Outline of a Non-Aristotelian Logic), in *Die Idee und Ihre Philosophischen Voraussetzungen, Volume 1,* F. Meiner, Ed. (Hamburg, 1959) 389 pp.

Gotthard Gunther, *Grundlagenstudien,* "Analog-Prinzip, Digitalmaschine und Mehrwertigkeit Grundlagenstudien" (Analogy, Digital Machines and Many-Valuedness), **1,** number 2, pp 41-50 (1960).

Gotthard Gunther, *Merkur,* "Schopfung, Reflexion und Geschichte" (Creation, Reflexion and History), **14,** pp. 628-650 (1960).

Gotthard Gunther, *Grundlagenstudien,* "Ein Vorbericht uber die Generalisierte Stellenwerttheorie der Mehrwertigen Logik" (Preliminary Report on the Generalized Place - value Theory of Many-Valued Logic), **1,** number 4, pp. 90-104 (1960).

Gotthard Gunther :"Cybernetic Ontology and Transjunctional Operations" in *Self-Organizing Systems,* M.C. Yovits and G.D. Goldstein, Eds. (Spartan Books, Washington, D.C., 1962) pp. 313-392.

Gotthard Gunther, *Soziale Wlt.,* "Logische Voraussetzungen und Philosophische Sprache in den Sozialwissenschaften" (the Tradition of Logic and the Concept of Trans-classical Rationality), **12,** pp 289-304 (1962).

Gotthard Gunther, *Allgemeen Nederlands Tijdschrift voor Wijsbegeerte en Psychologie,* "The Tradition of Logic and the Concept of Trans-classic Rationality", **54,** Number 4, pp. 194-200 (1962).

Gotthard Gunther, "Cybernetic Ontology and Trans-junctional Operations (Technical Report Number 4, Electrical Engineering Research Laboratory, Univeristy of Illinois, Urbana, 1962) 113 pp.

Gotthard Gunther, *Artorga*, "Hegel: Dialectic Theory and Symbolic Logic", **58**, pp. 1-7 (1963).

Gotthard Gunther, *Das Bewusstsein der Maschinen* (The Consciousness of Machines), (Agis Verlag, Baden-Baden, 1964, second, revised, and enlarged edition) 213 pp.

Gotthard Gunther, *Erstes Beiheft*, "Das Problem einer Formalisierung der transzendental-dialektischen Logic" (The Problem of a Formaligation of Transcendental-Dialectic), pp. 65-122 (1964; AFOSR 65-0698, AD617860).

Gotthard Gunther, "Information, Communication and Many-valued Logic", in *Memorias del XIII Congreso International de Filosofia*, **5**, Mexico, September 1964, pp 143-157 (AFOSR 64-2497).

Gotthard Gunther, *Zeitschr. f. Philos. Forschung*, "Zweiwertigkeit, Logische Paradoxie und Selbst-referierende Reflexion" (Two-valuedness, Logical Paradoxy and Self-referential Reflexion), **17**, Number 3, pp. 419-437 (1964; AFOSR 64-0343).

Gotthard Gunther, "Cybernetics and the Transition from Classical to Trans-Classical Logic" (BCL Report Number 3.0, University of Illinois, Urbana, 1964) 9 pp.

Gotthard Gunther, *Sprache im Technischen Zeitalter*, "Das Problem einer Trans-klassischen Logic" (The Problem of a Trans-classic logic), **16**, pp. 1287-1308 (1965).

Gotthard Gunther, "Uber Anschauung und Abstraktion" (Perception and Abstraction) in *Dialog des Abendlandes*, Physik und Philosophie, E. Heimendahl Ed. (List Verlag, Munchen, 1966) pp 199-207.

Gotthard Gunther, G.C. McVittie and Heinz Von Foerster, "Time", (BLC Report Number 3.1, University of Illinois, Urbana, 1966) 50 pp.

Gotthard Gunther, "Formal Logic, Totality and the Super-Additive Principle" (BLC Report number 3.3, University of Illinois, Urbana, 1966) 23 pp.

Gotthard Gunther, *Logik, Zeit, Emanation und Evolution* (Logic, Time, Emanation and Evolution), (Westdeutscher Verlag, Koln und Opladen, 1967) 96 pp.

Gotthard Gunther, "Time, Timeless Logic and Self-Referential Systems" in *Interdisciplinary Perspectives of Time*, R. Fischer, Ed. (New York Academy of Sciences, New York, 1967) pp. 396-406.

Gotthard Gunther and Heinz Von Foerster, "The Logical Structure of Evolution and Enanation", in *Interdisciplinary Perspectives of Time*, R. Fischer, Ed. (New York Academy of Sciences, New York, 1967) pp 874-891.

H.S.A. Na, Heinz von Foerster and Gotthard Gunther, "On Structural Analysis of Many Values Logic (Technical Report number 1, University of Illinois, Urbana, Department of Electrical Engineering, 1964; AFOSR 64-1381) 131 pp.

160

D.J. Schadach, "A Classification of Mappings Between Finite Sets and Some Applications" (BCL Report Number 2.2, University of Illinois, Urbana, 1967; AFOSR 67-1080, AD652234) 27 pp.

D.J. Schadach, "A System of Equivalence Relations and Generalized Arithmetic" (BCL Report Number 4.1, University of Illinois, Urbana, 1967) 34 pp.

6.5 APPLICATION OF WIENER CANONICAL EXPANSIONS TO PATTERN RECOGNITION

Donald B. Brick

Information Research Associates Inc., Cambridge, Massachusetts

Project Description

Research on Wiener canonical expansions of Bayes rules have suggested a powerful class of pattern recognizing automata applicable to a broad class of stochastic inputs. The automata, known as conditional probability computers, appear to avoid the major disadvantages of such computers, i.e., the tendency toward exponential growth as the number of learning samples increases. This effort will explore the usefulness and limits of this approach to the modeling of bionic systems and decision procedures. Limitations such as how broad a class or classes of processes can be accommodated by the procedure, the number of filter sections required as determined by the rate of convergence of the rate of convergence of the canonical series, stationarity and ergodicity requirements imposed on the inputs, and the self-adaptive or other intelligent properties the method provides will be investigated. The value of this method of structure synthesis will be compared with other methods such as power series or Edgeworth series expansion procedures.

Summary of Recent Research Results

A combined experimental (computer-simulation) /theoretical approach was initiated. The Hermite-Laguerre (Wiener canonical) expansion for the likelihood (or conditional probability) function was programmed and tested. A very careful and systematic investigation of the behavior of the expansion for the benchmark cases of white noise derived from a Wiener process (and the Wiener process, itself)* was undertaken since the response to these processes was completely understood from theory. Good convergence was obtained using a total of about 18 Laguerre and Hermite terms, and a complete understanding and profile of the behavior of the computer-simulation, including error analyses and approximations, have been gained. This has led to the formulation of a group of feature processors and preprocessors to aid in the convergence and accuracy of approximation and in recognition-and-intelligent system applications.

---

*It was known, a priori, that the expansion procedure is valid for such a process.

The theoretical effort has provided analytical backup and verification for the experimental effort and has, in addition, aimed at attaining a deep understanding of the processes involved and in defining the limits of applicability. To date, the following conclusions have been reached:

(1) All processes must be non-deterministic

(2) Certain continuity requirements must be met

(3) All processes derived via integrable operations on white noise may be handled

(4) The behavior of orthogonal increment processes is well defined

(5) Symmetrically distributed processes simplify the expansion.

Applications of the procedure have also been considered and enumerated.

Reports and Publications

Donald B. Brick, "On the Applicability of Wiener's Canonical Expansions", presented at the IEEE International Convention in March 1967 and it will appear in the January 1968 issue of the Transactions on Systems Science and Cybernetics (Technical Report No. 8, 20 March 1967; AD650193).

Donald B. Brick, "A Program for Adaptive Pattern Recognition" (Research Paper Dated 24 April 1967; to be published in a Book on Pattern Recognition, Laveen Kanal, Ed.).

Donald B. Brick, "On the Coefficients of a Wiener Canonical Expansion for the Likelihood Function of a Continuous Martingale" (Technical Report No. 10, 12 June 1967; AFOSR 67-1555, AD653305).

Donald B. Brick and J. Owen, "Pattern Recognition of Stochastic Processes (Revised)" (Technical Report No. 9, 23 June 1967; AD654199).

Donald B. Brick, H. Kashian, L. Petri and M.B. Robinson, "Digital Simulation of the Wiener Canonical Expansion" (Technical Report No. 7, July 1966; AD642712).

Donald B. Brick, Larry Petri and E.G. Henrichon, "Digital Simulation of the Wiener Canonical Expansion (Revised)" (Technical Report No. 13, 22 August 1967).

J. Owen, "Pattern Recognition of Stochastic Processes" (Technical Report No. 5, 7 July 1966; AD489190).

J. Owen, H. Kashian and Donald B. Brick, "An Investigation Into the Ensemble Properties of the Wiener Canonical Expansion" (Technical Report No. 4, 29 July 1966; AD488854).

R. Stout, "Fundamentals of the Wiener Canonical Expansion" (Technical Report No. 6, July 1966; AD641091).

6.6   **A FORMAL THEORY AND ALGORITHMS**
FOR INTELLIGENT ACTIVITIES

E. Mark Gold

Institute for Formal Studies,
Palo Alto, California

Reports Description

The ultimate objective is a machine which can do the following:
The machine communicates with a human by means of sensory
inputs and outputs and a distinguished reward input which the
human uses to indicate its approval of the machines responses. The
machine is to learn, by trial and error, to produce outputs which
evoke the human's approval. Such a machine can be achieved by
combining the following two operations: (1) A "Learner" observes
the inputs, outputs, and rewards of the machine, and learns to pre-
dict the reward. (2) A "generator" proposes output strategies
searching for one for which the learner predicts high likelihood of
reward. The current objective is a general learner which can be con-
nected to any sensory inputs and will learn to predict any judgement
which a human is asked to make. In particular, a learning theory is
sought which is applicable to linguistic analysis and pattern recog-
nition. It is to consist of (1) a model of the type of experiment in-
volved, (2) a schema for writing rules for analyzing sensory data and
making predictions, and (3) a practical algorithm for learning such
rules. Prior research constructed and investigated simplified models
of deterministic goal-seeking and learning situations, and formulated
the objectives of the learning theory which is sought presently.

Summary of Recent Research Results

1. A model of the type of experiment to which the desired learner
is to be applicable has been constructed.

2. Several related schema have been formulated for writing rules
for probabilistic analysis of sensory data and making probabilistic
predictions.

3. Two approaches are currently being pursued towards the con-
struction of the desired thinker: (3a) The learner continually uses

164

incoming data to construct new sensory analysis rules and refine old ones. For this purpose, principles for concept formation have been formulated fairly precisely. (3b) The results of all allowed sensory anal analyses are computed and averaged with respect to weights which are continually modified by incoming data. It has been proven that certain learners of this type can be guaranteed to learn, in a certain sense, any of the allowed sensory analysis.

Reports and Publications

E. Mark Gold, *Journal of Symbolic Logic*, "Limiting Recursion", **30**,pp. 28-48 (1965).
E. Mark Gold, *Information and Control*, "Language Identification in the Limit", **10**, pp. 447-474 (1967).

6.7 THE EXISTENCE AND USE OF ALGORITHMS

Martin D. Davis and Hilary Putnam

New York University, New York, New York

Project Description

Modern developments in mathematical logic are emerging as the tools or prescriptions for the explicit representation of procedures generally labeled "intelligent" activities. Efforts to prove mathematical theorems by logical algorithms which are amenable to digital computer processing constitute one approach toward the methodology for the reduction of "intelligent" activities to machine recognizable forms. These studies propose to explore proof procedures a) on a problem in the elementary theory of groups; b) on the problem of identity; c) on problems in number theory from developments of finitely axiomatized systems or infinitely axiomatized systems with a finite set of axioms and induction schema; and d) on systems of notation for transfinite ordinal numbers. The study of semi-computable (recursively enumerable) sets is also planned in an attempt to extend earlier results (Annals of Mathematics 74:425-436, 1961) relevant to number theory and automata theory.

Summary of Recent Research Results

Simplification of Takeuti's ordinal diagrams have been obtained leading towards comparison with Schutte's $\Sigma(m)$. Non-existence of non-trivial integer solutions to $9(u^2 + 7v^2)^2 - 7(r^2 + 7s^2)^2 = 2$ has been shown to imply that every semi-computable set is Diophantine. The word problem for semi-groups with one relation has been solved in certain special cases. Martin-Lof random sequences have been found in $\Sigma_2 \cap \pi_2$. Improvements have been obtained in characterizations of ordinal numbers by using the quantifier: "there are infinitely many".

Reports and Publications

M.M. Zucherman "Finite Versions of the Axiom of Choice" (Technical Report ALG - 1, June, 1967; AFOSR 67-1520, AD655086).

H. Levitz, *Commentarii Mathematici Helvitici*, "Uber die Finslerschen Hohern Arithmetischen Operationen", **41**, pp. 17-30 (1967).

166

6.8   THEOREM PROVING BY COMPUTER

Bruce Kallick

Northwestern University, Evanston, Illinois

Project Description

Problem-solving techniques normally used by man need not be the most efficient procedures for machines. Some of the possibilities offered by high-speed computers for non-numeric information processing are being explored in studies concerned with devising algorithmic procedures capable of computer implementation for proving mathematical theorems. For machine recognition of the notion of proof, a two-part formulation is required: a translation from natural language into a formal language, and formal rules defining what is meant by a proof. Exhaustive enumeration in predicate calculus proof procedures is a crucial weakness of the scheme. This effort seeks solutions to the problem via the Herbrand theorem by expressing the denial of the theorem as a prenex formula of predicate calculus and generating an inconsistent set of instances of this formula. Reformulation of a generalization of a decision procedure in terms of conjunctive normal form matrices rather than disjunctive normal form matrices has eliminated the need of special techniques for amplifying the matrix.

Summary of Recent Research Results

During the last year a proof procedure has been developed that incorporates the most attractive features of two of our most efficient algorithms. A small scale version of the procedure has been programmed (on a CDC 3400) to process formulas of the $E_4A_2E_2$ prefix class. One novel feature of the program is that it is designed to deal with formulas of predicate calculus with equality; the prefix class was chosen to include several examples that have been proposed as test cases for such programs. A second program deals with formulas of the $A_1 E_1 A_1$ class with monadic and dyadic predicates. Experimentation with this highly restricted class of formulas will continue due to the simplicity of this class of formulas and our interest in finding decidable subclasses.

167

Reports and Publications

Bruce Kallick, "Theorem-Proving by Computer" (Final Report, Jan. 1965; AFOSR 65-0338, AD611815);

Bruce Kallick, "Theorem Proving by Computer" (Final Report, Jan. 1966; AFOSR 66-0181, AD628319);

168

6.9   **LOGIC AND RATIONAL PROCESSING**
IN SCIENTIFIC DISCOVERY

Norwood Russel Hanson* and Michael Scriven

PEC, Inc., Athens, Georgia

Project Description

Attempts to analyze scientific inference and describe the logical
processes of scientific discovery, explanation, and prediction are
usually ex post facto reappraisals of finished reports and do not
take into consideration the rational processing of information in-
volved in the resolution of the perplexities of anomalous situations
preceding the final results. This study seeks to provide, in rigorously
analytic terms, a description of the rational processes involved in
going from an initial scientific inquiry to a result, with emphasis on
the complete chain of reasoning, selection and modification of
criteria, and evaluations of complex sets of data which are the bases
for achievement of the result. Relationships between explanation and
prediction will be examined in detail. This explication of the structure
of discovery should enhance understanding of the nature of scientific
inquiry and the logic of problem solving and facilitate the recognition
of intelligent and creative activities.

Summary of Recent Research Results

In April 1967, Professor Norwood Russell Hanson died in an
aircraft accident. Professor Michael Scriven of the Philosophy
Department, University of California at Berkeley is reviewing
Professor Hanson's unpublished papers in an attempt to
convert this material to publishable form.

---

*Deceased.

6.10 TRANSFORMATION-GENERATING PRINCIPLES IN BIOLOGICAL SYSTEMS

John M. Myers

Raytheon Company, Waltham, Massachusetts

Project Description

The value of biological systems as paradigms of adaptive, control, and "intelligent" or living-system-like devices is predicated on an understanding of their structures and the principles by which they function. To date, models of neurons and neural nets and adaptive devices fall far short of explaining or approximating the performance of biological systems. An hypothesis of Lettvin is that each neural response pattern is unique, dependent in part on its environment, and its performance is not amenable to truth-table logic. The rules by which neurons interact are viewed as too incomplete and contingent to be reducible to logical formalisms. This effort suggests an analogy between biological systems and complex devices and organizational systems and assumes that both biological and engineering systems depend on an interplay between the performance of a germinal cell and a generating principle, the generating principle providing for complexity by the elaboration of simple incremental elements. The effort seeks to develop a method of comparing cells and patterns of interaction among cells that is free of the assumption that cells can be analyzed in detail. An attempt will be made to characterize selectio processes that results in variations in behavior, dominance of some cells over others, etc. A language for describing such systems would be applicable to the construction of perceptive devices.

Summary of Recent Research Results

This research effort was concluded after five months of its inception since the principal investigator accepted a government position with the Office of the Secretary of Defense. The final report on five months of research does not contain "results" in the usual sense, but is an account of the inquiries and reflections of a physicist and applied mathematician who explored problems in biology from the viewpoint that the development of information processing systems of organisms must occur in a sufficiently rule-like way to be mathematically describable. The report considers: (1) bases for the supposition of transformation generating principles that are rooted both in evolutionary and developmental phenomena; (2) a comparison of beginnings of results in artificial intelligence and

170

results of research on relevant biological structure; and (3) efforts
to devise new kinds of mathematics for biology. Analogies are dis-
cussed between electronic and nervous circuits with the observation
that "the analogy between the brain and a very fancy circuit is an
analogy between a brain which we do not understand and a 'circuit'
that is completely outside our experience with circuits and which
we do not understand either." Considerations of perception and
context lead to the statement that "an organism can be taken as the
expression of a point of view, and a point of view can be taken as a
partitioning of the universe into classes of equivalence of response."
Thus, a characteristic feature of a mathematical representation of
structure would come from the credibility that a point of view
attributes to other points of view. Though consideration of relations
between points of view led back to those that stem from sets, a set
may be an expression of a point of view that is yet, in the biologi-
cal sense, to be described mathematically.

Reports and Publications

John M. Myers, "In Search of Mathematics for Biology" (Final Tech-
nical Report, S-932 April 1967; AFOSR 67-0893, AD 650404).

SYSTEMS OF NATURAL DEDUCTION WITHOUT ESSENTIAL RESTRICTIONS ON VARIABLES

Kurt Bing

Rensselaer Polytechnic Institute, Troy, New York

Project Description

Research that attempts to simplify formal representations
of logical systems has implications for new logical designs for
machines and for methods of manipulating complex information
structures. This effort is for the study of systems of natural deduction
in which two quantifier rules are replaced by essentially equivalent
ones that do not require the usual somewhat cumbersome restrictions
on the free variable and result in a more convenient formalism.
The system is equivalent to the Kleene Hilbert-type system for the
predicate calculus modified by excluding variation of variables. The
approach will be applied to linear systems of natural deduction such
as those of Quine or Suppes. The resulting linear system is expected
to have the subsystem and separation properties of the Gentzen
and the Kleene systems. A device for combining deductions beyond
the rules of existing linear systems may be required.

Summary of Recent Research Results

A linear system of natural deduction has been devised which
incorporates the new quantifier rules previously described. No
variables are flagged. The system is complete as to consequence and
has the anticipated subsystem and separation properties. A preliminary
analysis indicates that the intuitionistic subsystem, with possible
slight alterations, has the subformula property for suitable deductions.

Reports and Publications

Kurt Bing, "A System of Natural Deduction with few Restrictions
on Variables", in *Proceedings Hughes Aircraft Company Symposium
on Logic, Computability, and Automata*, Rome Air Development
Center, Rome, New York, 1965 (to be published).

Kurt Bing, "A Linear System of Natural Deduction with Few Restrictions
on Variables", *Third International Congress for Logic, Methodology
and Philosophy of Science*, Abstracts of Papers, Amsterdam,
Netherlands, 1967.

6.12   LOGIC FOR SELF-ADAPTING AUTOMATA AND UNIVERSAL ELEMENTS

John F. Randolph

University of Rochester, Rochester, New York

Project Description

Automata capable of adaption and self-repair may be the solution to the reliability problem of complex electronic equipment to be used in future space efforts. This research effort studies the design and characteristics of logic elements that are capable of adaption or reorganization. It will include research on the characteristics of "Universal Logic Elements" (single elements capable of performing all logic functions). Particular attention will be directed towards definitizing functionally optimum network configurations, yet minimizing error probability. This work will include a) investigations of network configurations which tolerate the greatest probability of error, and b) studies of the stochastic functions which describe self-organizing networks.

Summary of Recent Research Results

Since adaptive characteristics are exhibited in all living organisms and have been studied most extensively in the human, the analysis of adaptation began by axiomatizing an accepted verbal theory under the general title, "PSYCHO-LOGICS" (human behavior in the (Organism x Environment) field). Three papers have been submitted for publication. They are: 1. "A Logical Calculus of Behavior"; 2. "The Structure of Unimpaired Human Behavior"; and 3. "The Structure of Empathy". A forth paper entitled, "The Structure of Impaired Human Behavior" is nearing completion.

This Research was done in close collaboration with C.A. Hilgartner, M.D. who specialized in Psychotherapy but who currently is engaged in biochemical research on gene structure. The services of a neurophysiologist and an electrical engineer have been enlisted to work on methods of effecting adaptation, since outward manifestations of behavior are on a firm (set theoretic/ symbolic logic) foundation. Hopefully, these efforts will lead to a realistic design of electro-chemical adaptive systems.

6.13 STUDIES IN LOGIC ON INFORMATION CAPACITY AND COMMAND HIERARCHIES

W.H. Hanson and Marius Cohn

Sperry Rand Corporation, St. Paul, Minnesota

Project Description

Studies in logic are being pursued to gain an understanding of possibilities and limitations, both in the manipulation of information and in the specification of structures (algorithms) for such manipulation. This effort is examining the applicability of logic for two purposes: (a) to obtain a precise, quantitative definition of (information capacity of a logic expression) (i.e., Boolean or similar algebraic formulas) and theorems relating the definition to applications, and (b) the development of a formal model of complex command hierarchies which includes a decision procedure that can be applied to logics containing many command operators. A definition of information capacity has been established for propositional calculus expressions together with an algorithmic procedure for its computation for any such expression. Information capacity will be examined as a measure for the complexity of a function which should be known precisely in determining the merit of a logic design. Given a command hierarchy, an adequate command logic is envisaged which would provide a formal language usable by members of the hierarchy for internal communications and by observers for describing and evaluating the operations of the hierarchy. The decision procedure is expected to incorporate modifications of existing procedures for alethic modal logics.

Summary of Recent Research Results

While no satisfactory measure of absolute complexity of a logical network (in terms of information capacity) has yet been found, a possible means for comparison of the informational capacities of networks has been developed and is being studied. A most difficult problem arises in attempting to define the state of *zero* initial information.

A formal logical structure for command systems has been developed and is now being examined for its applicability to real-world situations. This work is described in (1).

Reports and Publications

W.H. Hanson, *Logique et Analyse*, "A Logic of Commands" (December, 1966).

174

6.14   **RESEARCH IN AUTOMATA THEORY
AND ITS MODELING APPLICATIONS**

Michael A. Arbib

Stanford University, Stanford Electronics Labs
Stanford, California

Project Description

Automata Theory is a generic term that refers to numerous
current inquiries into mathematical structures that represent
various characteristics of information processing and may yield
new logical designs for machines. This effort will examine
several structures and relate them to information processing
pnenomena. Extensions of descriptions of finite-state machines
are sought either through a more natural algebraic structure than
the semigroup or through modification of semigroup theory to
provide for feedback elegantly analogous to a description for
loop-free decomposition An effort will be made to extend
common bases of automata and control theory. Automata-
theoretic methods will be applied to self-reproducing models
that have relevance to the design of parallel computers.
Relationships will also be studied between automata theory and
psychological processes disclosed by information-processing
languages.

Summary of Recent Research Results

1. *Theory of Automata*

Three papers have resulted from this investigation of the structure
of finite automata using algebraic methods. (1) developed the notion
of machine-diagram chasing, and used this technique to give a
complete analysis of the relation between decomposition of
machines and extensions of semigroups; (2) used the machine-
diagram chasing technique to give a strengthened version of the
Generalized Embedding Theorem and extend the Krohn-Rhodes
theory of automaton complexity, to explicate the manner in
which wreath products increase machine complexity. Structure-
preserving transformations of the state-spaces of machines,
with a strong analysis of parallel compositions of machines, have
been presented on (3).

We are continuing our study of stochastic automata.(4), and
have made progress in elucidating problems of feedback in

automata, obtaining a significant generalization of Winograd's results on time as a measure of complexity of computation (5). We believe that this approach complements the Krohn-Rhodes approach to computational complexity, and we plan to explore this complementarity to build up a deeper picture of complexity of automata.

2. *Modelling Applications*

We have been able to apply our study of self-reproducing automata to obtain new insights into problems of embroyological development (6). We are new introducing more parallel processing into our model, and couducting intensive discussions with biologists to increase its relevance. Our self-reproducing models have great relevance to the design of highly parallel computers, and we are now turning some of our research to this aspect of computer design.

In collaboration with Dr. Roy Kahn, a Boston child psychologist, we have produced a cybernetic discussion of childhood information-processing which sheds some light on the processes of infantile autism and childhood schizophrenia (7).

In collaboration with Dr. Gene Franklin of Stanford and Dr. Nils Nilsson of Stanford Research Institute, we have produced a pilot study (8) of the role which computer science and system theory can play in elucidating the cerebellar function, and we new hope to analyze certain aspects of cerebellar disfunction in terms of our automata-theoretic study of complexity of computation.

Reports and Publications

Michael A. Arbib, "Automaton Decomposition and Semigroup Extensions", in *Algebraic Theory of Machines, Languages and Semigroups*, M.A. Arbib, Ed., (Academic Press, to be published).

Michael A. Arbib, J.L. Rhodes and B. Tilson, "Complexity and Group-Complexity of Finite-State Machines and Finite Semigroups", in *Algebraic Theory of Machines, Languages and Semigroups*, M.A. Arbib, Ed., (Academic Press, to be published).

Michael A. Arbib, *Inf. Control* "Automaton Automorphisms", to be published.

Michael A. Arbib, *Ann. Math. Statish.*, "Realizations of Stochastic Automata", 38, pp. 927-933 (1967).

P.M. Spira and M. A. Arbib, "Computation Times for Finite Functions, Automata and Groups", in *IEEE 8th Annual Symposium on Switching and Automata Theory*, Convention Record (1967).

176

M.A. Arbib, *J. Theoret. Biol.* "Automata Theory and Development, Part I", **14**, pp. 131-156 (1967).

R.M. Kahn and M.A. Arbib, "Cybernetics and Mental Development", to be presented at the 45th Annual Meeting, American Orthopsychiatric Association, March, 1968.

M.A. Arbib, G.F. Franklin and N. Nilsson, "Some Ideas on Information Processing in the Cerebellum", to appear in the proceedings of the Summer School on Mathematical Models of Neuronic Networks, Ravello, Italy, June 4-19, 1967.

6.15   A BRAIN MODEL USING MULTI-VALUED LOGIC

Alfred L. Stern

Wayne State University, Detroit, Michigan

Project Description

Considerable study is being made of biological nervous systems to determine the mechanisms by which these systems function and perform on incomplete information. Models of the activities of these systems have been constructed in the framework of classical two-valued logic which is the theoretical basis for most current digital computers. These models, however, involve either an impractically large number of connection diagrams or are not representative of the biological systems. This effort was directed toward finding a proper model for the information handling capability of the central nervous system with respect to its perceptual and storage capacity and its information handling techniques in terms of multivalued-logic representations. Ternary logic was used in a first approximation. A simulated ternary environment and the generation of ternary objects in it have been programmed. Simulations were also programmed that describe ternary subsystems for senory input, transmission, and memory. Results were extrapolated to a quaternary logic. Based on results with ternary and quaternary logics, the order of complexity will be determined for a system using multivalued elements (with operations of the same logical order) which would be needed for a proper model of the central nervous system.

Summary of Research Results

A series of distinctions in the uses of logical frameworks led to a proof of the non-reductive character of n-valued logics. An outline of these distinctions and the resulting proof appears in the *1964 Proceedings of the International Congress of Logic, Methodology, and Philosophy of Science.* An investigation of the representative uses of a ternary logic shows an increase in information capacity without interpretive schema over diadic representation of the expected magnitude but a much greater increase when interpretive frameworks would be naturally adopted as in decision procedures. This would be true whether or not the increased functional richness of the ternary logic is employed. The model finally suggested utilized quaternary logic for representation of input in the discriminatory network while a ternary logic was used for decision and judgemental procedures. The basis of this model can be extrapolated to

178

show an economic use of multivalued logic in a system such as the central nervous system. It would be likely that different physical representations of different end-valued logics would be found with different biological concomitants in the perceptual, information, control, and decision functions of the central nervous system.

SYMBOLIC LOGIC (SET THEORY, PROOF THEORY, ABSTRACT QUANTIFICATION THEORY)

Raymond M. Smullyan

Yeshiva University, New York, New York

Project Description

Research in symbolic logic in recent years has been directed
toward the development of methods of proof similar to those of
ordinary mathematical practice and amenable to digital computer
processing. Some of the newer methods are believed interrelatable in
previously unsuspected ways which may be useful for such problems
as combinatorial problems, particularly to the theory of graphs.
This effort is for an investigation of applications of the theory of
computable (recursive) functions to computing machines and finite
automata. A unifying principle have been demonstrated. Modifications
of this principle for intuitionistic logic and recursive function theory
are being explored. Unifying principles are also being sought in set
theory. Improved symbolic methods of reasoning are expected to
have ultimate application in the exact sciences.

Summary of Recent Research Results

Some research results since January 1966 are as follows:

1. Demonstration of a unifying principle in set theory - the
   double superinduction principle - which reduces to a common
   construction proofs of the well ordering theorem, the trans-
   finite recursion theorem and several results on ordinal
   numbers. This also leads to another equivalent of the axiom
   of choice.

2. An improved version of the tableau method for Quantification
   Theory which, unlike Beth Tableaux, use only one tree instead
   of two. This leads to considerable combinatorial simplifications
   in the metatheory e.g. in the proof of Gentzen's Hauptsatz.
   In connection with this tableau method, stronger versions of
   the Completeness Theorem resulted with applications to new
   systems of Linear Reasoning which avoid appeal to prenex
   normal form and which can be modified for Intuitionistic
   Logic.

180

3. Cut-free Natural Deduction Systems and their relation to tableaux. The significant relationship is really graph-theoretic rather than logistic. The graph-theoretic aspects of Natural Deduction and Tableaux were abstracted into "nest *structures*" and *"trees"* - the translation process involved is a special case of a theorem obtained on isomorphism between trees and nest-structures.

Reports and Publications

Raymond M. Smullyan, *Journal of the Mathematical Society of Japan*, "Pseudo-Uniform Reducibility", **15**, Number 2, pp. 129-133 (April 1963; AFOSR 1147; AD423135).

Raymond M. Smullyan, "A Unifying Principle in Quantification Theory, in *Proceedings of the National Academy of Science*, **49**, Number 6, pp. 828-832 (June 1963; AFOSR 1146; AD423138).

Raymond M. Smullyan, *Trans, Am. Math. Soc.*, "Creativity and Effective Inseparability", **109**, Number 1, pp 135-145 (October, 1963; AFOSR 64-0231, AD432519).

Raymond M. Smullyan, *J. Symbolic Logic, "Trees and Nest Structures"*, **31**, Number 3, pp. 303-321 (September 1966; AFOSR 67-1504, AD654613).

Raymond M. Smullyan, *J. Symbolic Logic*, "Finite Nest Structures and Propositional Logic", **31**, Number 3 pp 322-324 (September 1966; AFOSR 67-1525, AD654755).

Raymond M. Smullyan, *Trans. NY Acad. Sci.*, "On Transfinite Recursion", pp 175-185 (December 1966; AFOSR 67-1524, AD654754).

Section 7

INDEX TO INSTITUTIONS

Section 8

INDEX TO PRINCIPAL INVESTIGATORS

188

Section 9

INDEX TO ARPA SUPPORTED PROJECTS MONITORED BY THE DIRECTORATE OF INFORMATION SCIENCES

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
Air Force Office of Scientific Research Office of Aerospace Research, USAF Arlington, Virginia 22209	Unclassified
	2b. GROUP

3. REPORT TITLE

INFORMATION SCIENCES - 1967

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific   Special

5. AUTHOR(S) (First name, middle initial, last name)

Harold Wooster, Eliot Sohmer (Editors)

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
January 1968		

8a. CONTRACT OR GRANT NO. In-house	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO.	
c. 61445014	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d. 681304	AFOSR 68-0006

10. DISTRIBUTION STATEMENT

1. Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
TECH, OTHER	Air Force Office of Scientific Research Directorate of Information Sciences (SRI) Arlington, Virginia 22209

13. ABSTRACT

This is the fourth in a series of annual reports of the basic research program of the Directorate of Information Sciences, Air Force Office of Scientific Research. Summaries of projects current during fiscal year 1967 (1 July 1966 through 30 June 1967) are given within six categories: (a) information systems research, (b) information identification and classification, (c) transmission of information, (d) adaptive and self-organizing systems, (e) language and linguistics research, and (f) theoretical foundations of information sciences. Brief comments are made about each category. Indexes to institutions, principal investigators, and ARPA supported efforts monitored by the directorate staff are included.

DD FORM 1473, 1 NOV 65

190

Information sciences
Programming (computers)
Bibliographies
Digital computers
Computers
Scientific research
Computer logic
Classification
Documentation
Computer storage devices
Information retrieval
Language
Speech
Hearing
Nervous system
Mathematical logic
Pattern recognition
Speech transmission
Machine translation
Set theory
Cybernetics
Social communication
Nerve cells
Psychoacoustics
Bionics
Artificial intelligence
Perception (psychology)
Auditory perception
Autonomic nervous system
Speech recognition
Man-machine systems
Libraries
Automata
Information theory
Linguistics
Semantics
Syntax
Computational linguistics
Context free grammars
Subject indexing